Benford's law as mass movement detector in seismic signals

Qi Zhou[1], Hui Tang[2], Jens Martin Turowski[2], Jean Braun[2], Michael Dietze[3], Fabian Walter[4], Ci-Jian Yang[2], and Sophie Lagarde[2]

[1]Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences
[2]Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences
[3]Georg-August-University Göttingen
[4]Swiss Federal Institute for Forest, Snow and Landscape Research (WSL)

July 20, 2023

Abstract

Seismic instruments placed outside of spatially extensive hazard zones can be used to rapidly sense a range of mass movements. However, it remains challenging to automatically detect specific events of interest. Benford's law, which states that first non-zero-digit of given datasets follow a specific probability distribution, can provide a computationally cheap approach to identifying anomalies in large datasets and potentially be used for event detection. Here, we select raw seismic signals to derive the first-digit distribution. The seismic signals generated by debris flows, landslides, lahars, and glacier-lake-outburst floods follow Benford's law, while those generated by ambient noise, rockfalls, and bedload transports do not. Focusing on debris flows, our Benford's-law-based detector is comparable to an existing random forest method for the Illgraben, Switzerland, but requires only single station data and three non-dimensional parameters. We suggest this computationally cheap, novel technique offers an alternative for event recognition and potentially for real-time warnings.

Benford's law as mass movement detector in seismic signals

Qi Zhou[1, 2†], Hui Tang[1], Jens M. Turowski[1], Jean Braun[1, 2], Michael Dietze[3, 1], Fabian Walter[4], Ci-Jian Yang[1, 5], Sophie Lagarde[1, 2]

[1]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam,

[2]Institute of Geosciences, University of Potsdam, Potsdam,

[3]Faculty of Geoscience and Geography, Georg-August-Universität Göttingen, Göttingen,

[4]Swiss Federal Institute for Forest, Snow and Landscape Research, Zürich,

[5]Department of Geography, National Taiwan University, Taipei,

Corresponding author: Qi Zhou (qi.zhou@gfz-potsdam.de, F427, Telegrafenberg, 14473 Potsdam, Germany)

Key Points:

- The first-digit distribution of seismic signals generated by high-energy mass movements and fluvial processes follows Benford's law
- When Benford's law appears, raw seismic signals tend to increase exponentially and converge to a power law distribution with exponent one
- A computationally cheap and novel detector based on Benford’s law is developed for debris-flow events

Abstract

Seismic instruments placed outside of spatially extensive hazard zones can be used to rapidly sense a range of mass movements. However, it remains challenging to automatically detect specific events of interest. Benford's law, which states that first non-zero-digit of given datasets follow a specific probability distribution, can provide a computationally cheap approach to identifying anomalies in large datasets and potentially be used for event detection. Here, we select raw seismic signals to derive the first-digit distribution. The seismic signals generated by debris flows, landslides, lahars, and glacier-lake-outburst floods follow Benford's law, while those generated by ambient noise, rockfalls, and bedload transports do not. Focusing on debris flows, our Benford's-law-based detector is comparable to an existing random forest method for the Illgraben, Switzerland, but requires only single station data and three non-dimensional parameters. We suggest this computationally cheap, novel technique offers an alternative for event recognition and potentially for real-time warnings.

Plain Language Summary

Natural hazards, such as debris flows and landslides, pose a significant threat to the exposed communities. Seismic instruments as seen as effective tools for detecting these hazardous processes and may be used in early warning systems. However, the difficulty lies in identifying the events of interest concisely and objectively. Our study explores Benford's law, a probability distribution of the first-non-zero digit. We collected seismic data generated by various hazard events and compared the observed first-digit distribution with their agreement with Benford's law. We found seismic signals of high-energy mass movements follow Benford’s law during the running phase, while ambient noise and other small mass movements do not. In order to explain why Benford’s law is followed, we argue that raw signals increase exponentially and fit a power law distribution with exponent as one. Our detector, based on Benford's law and designed for debris flow, which is a computationally cheap and novel model, performs similar to a machine learning algorithm previously used in the study site. Our work illustrates a new approach to detecting events and designing warning systems, which can be used in different regions.

Keywords Environmental seismology, mass movement, Benford’s law, event detector, debris flow, early warning system.

1 Introduction

Mass movements (e.g., landslide and debris flow) and extreme fluvial processes (e.g., flash floods and glacier-lake-outburst flood) are of significant concerns in populated areas, as they can cause huge loss of life and damage to civil infrastructure each year (Holub & Hübl, 2008; Merz et al., 2021; Regmi et al., 2015). Classification criteria for mass movements may vary depending on the focus of interest (Coussot & Meunier, 1996; Nemčok et al., 1972). Yet, the most widespread and destructive mass movements are generally considered to be debris flows, landslides, and rockslides (Dowling & Santi, 2014; Froude & Petley, 2018). Despite extensive efforts to mitigate their hazard through risk assessment and structural measures (Dai et al., 2002; Fuchs et al., 2007; Huebl & Fiebiger, 2015), the intricate geological conditions and dynamic processes of mass movements frequently pose challenges in preventing property damage and fatalities (Fan et al., 2019; Kean et al., 2019; Tiwari et al., 2022).

Early warning systems are an established approach to mitigating the impact of mass movements (Badoux et al., 2009; Guzzetti et al., 2020; Hürlimann et al., 2019). For example, systems based on measured rainfall intensity and predefined thresholds for triggering alarms (Baum & Godt, 2010; Marra et al., 2016) are among the most popular warning approaches. However, maintaining rain gauges and obtaining accurate rainfall intensity data in real-time is challenging for the operation of a warning system, especially for catchments with large elevation differences. Inaccurate measurements and uncertainty in data interpolation lead to significant errors in rainfall thresholds (Nikolopoulos et al., 2015). In addition, due to the variability in geological and hydrological conditions, empirical thresholds for triggering debris flows and landslides are not transferable between catchments (Gregoretti et al., 2016; Wilson & Wieczorek, 1995). Detecting specific events of interest from time-series signals is essential for releasing a warning. Force plates, radar, laser, and video cameras are the most common sensors used for monitoring in early warning systems (Comiti et al., 2014; McArdell et al., 2007). However, some of these devices require a high-power supply and regular maintenance, and can be easily destroyed by the hazard processes itself.

Continuous seismic and acoustic signals offer a new way to monitor mass movements with high temporal resolution (Le Breton et al., 2021; Burtin et al., 2016; Cook & Dietze, 2022; Farin et al., 2019; Schimmel et al., 2013). The instruments can be installed outside the zones affected by the hazard and are thus in lesser danger of being destroyed. An array of seismic stations can help to detect and locate extreme, high-energy events on a regional scale (Cook et al., 2021; Ekström & Stark, 2013; Hammer et al., 2012). However, a seismic station records all ground vibration signals within its bandwidth, blending events of interest and those considered as noise. Current seismology-based detectors of mass movements and fluvial processes, such as seismic attributes-based methods (Dietze et al., 2022; Govi et al., 1993; Schimmel & Hübl, 2016; Wei & Liu, 2020), short-term average to long-term average ratio (Coviello et al., 2019), random forests (Hibert et al., 2019; Provost et al., 2017), and hidden Markov models (Dammeier et al., 2016; Hammer et al., 2012) require numerous waveform, spectral, network features or parameters to be fed into the model to identify events. In addition, collecting and labeling the data to parameterize or train such a model is time-consuming and requires experience. Applying these existing approaches to other sites requires re-training the model or calibrating the parameters; worse, often no historical data are available for most new sites to do this. Before warning systems can be constructed, implemented, and promoted, a convenient and portable approach to event detection must be found. Compared to ambient noise and signals not associated with extreme events in a natural environment, the temporal occurrence probability of mass movements is relatively low. Therefore, detecting debris flows and other mass movements in seismic time-series signals can be treated as an anomaly detection.

The Newcomb–Benford law (BL) or the first-digit law, which is widely used in fraud and data quality detection, is a probability distribution of the first digit of a dataset (Castañeda, 2011; Cho & Gaines, 2007; Ley, 1996). Newcomb (1881) stated that the probability of occurrence of the first digits is such that the mantissae of their logarithms are equally probable:

$$P(d) = \log_{10}(1+d^{-1}) \tag{1}$$

where $P(d)$ is the theoretical probability of the first none zero digits, d={1, 2, …, 9}. For example, -0.01 and 100 share one as the same first digit with a likelihood of 0.301. Frank Benford rediscovered this relationship, tested it with twenty different datasets. It was later named after him as Benford's law (Benford, 1938). BL has been used to several fields of the geosciences, such as in studying the homogeneity of natural hazard datasets and anomalies

(Geyer & Martí, 2012; Joannes-Boyau et al., 2015). Earthquakes and Mars quakes were detected in seismic signals with BL (Díaz et al., 2015; Sambridge et al., 2010; Sun & Tkalčić, 2022). Due to the dimensionless and low computational cost of BL, it has the potential to be used to identify mass movements in seismic data at different catchments, perhaps even as detector in data loggers.

In this study, we compiled seismic data generated by various mass movements and fluvial processes, calculated the first-digit distribution of seismic signals, and investigated which processes or periods follow the BL. We explain why BL appears in seismic signals generated by some of the processes and not by others. Finally, we present a BL-based event detector for debris flows and compare its performance with a previously developed random forest model (Chmiel et al., 2021) for the same seismic network. This work shows a novel approach for detecting high-energy mass movements and the potential for establishing a real-time warning system using BL.

2 Data Source and Event Catalog

2.1 Study Site and Data Source

The Illgraben catchment near the village of Leuk, southwest Switzerland (Figure 1a) is one of the most active debris flow catchments in the Alps. It covers an area of about 9.5 km^2 and extends from the Rhône River at 610 m to the Illhorn Mountain, peaking at 2716 m (Badoux et al., 2009). The annual rainfall is concentrated from May to October, and the Illgraben catchment roughly experiences three to five debris flows and several floods each year, mainly triggered by short-duration convective storms (McArdell et al., 2007). To mitigate the risk of debris flows and floods, a warning system has been implemented at the Illgraben that triggers an alarm when the impulse of in-torrent ground vibration sensors exceeds empirically determined thresholds (Badoux et al., 2009).

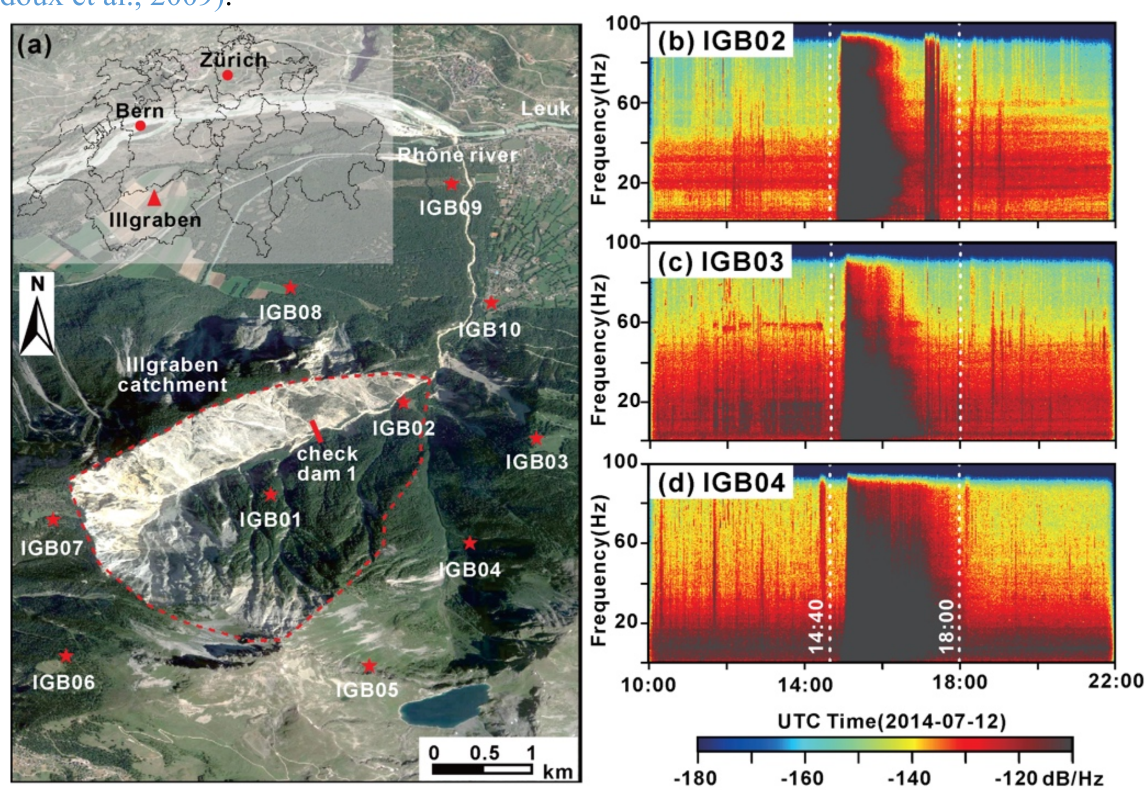


Figure 1 Study area and debris flow at Illgraben catchment. (a) Location and distribution of seismic stations (red stars, Table S1 for details). (b) to (d) are the spectrograms of the vertical component for a debris flow event between 14:40 and 18:00, 12th July 2014.

Considering that seismic data are not available for all stations for the whole year and the complexity of signals from stations far from the spatially propagating event, we mainly selected the IGB02 station for this study (same location as ILL02/ILL2 deployed by Swiss Federal Institute for Forest, Snow and Landscape Research WSL), which is closest to the channel and far from the nearby residential area of Leuk.

2.2 Event Catalog

To calculate the first-digit distribution of seismic signals and quantify which processes or periods follow BL, we examined 24 debris flows (one of which may be a flood event) that occurred 2013-2014 and 21 debris flows that occurred 2017-2019 (Tables S2-S3) in the Illgraben catchment. For the 2013-2014 debris flows, ten of the 24 events were recorded by local warning systems (WSL events), and we manually labeled an additional 14 debris flows based on the event duration of waveforms and the 1-50 Hz features of the spectrogram (GFZ events, Text S1). One example of the debris flow that occurred on 12th July 2014, with the WSL label and high signal-to-noise ratio SNR (about 20 based on IGB02), is shown in Figures 1b-1d.

To complement the data with events from other locations and instruments for calculating the first-digit distribution, we added seismic signals from other mass movements and fluvial processes (Table S4), such as a 2013 rockfall event in Illgraben, Switzerland (Burtin et al., 2016), a 2015 rockfall event in Lauterbrunnen, Switzerland (Dietze et al., 2017), a 2014 landslide in Askja, Iceland (Schöpa et al., 2018), a 2015 hurricane-induced lahar in Volcán de Colima, Mexico (Capra et al., 2018), a 2016 glacial-lake-outburst flood GLOF in Bhotekoshi, Nepal (Cook et al., 2018), and a 2019 bedload transport event in Liwu catchment, Hualien.

3 Methods

3.1 Data Preparation

Processing seismic signals using demeaning, detrending, filtering, or deconvolution may alter the first-digit distribution and obscure the difference between BL in ambient noise and the event phase (Figure S1). Therefore, we use the raw vertical-component seismograms (units are counts) to calculate the first-digit distribution and check whether the observed distribution adheres to BL. We choose a one-minute moving window (no overlap) to calculate the probability distribution of digits one to nine to avoid statistical errors associated with a small dataset. The number of data points (n) for each window is equal to the moving window length (W_L, units are seconds) multiplied by the sampling frequency (f_S, units are Hertz, Table S1 for details):

$$n = W_L * f_S \tag{2}$$

For each window, data points with a raw amplitude equal to zero are discarded.

3.2 First-digit Distribution and Benford's Law

We used two established statistical methods, the Chi-squared test (Geyer & Martí, 2012; Patefield, 1981) and the Kolmogorov-Smirnov test (Kaiser, 2019; Feller, 1948), to validate whether the observed first-digit distribution follows BL. The hypothesis is that the frequency of the observed first digits is not distinct from the theoretical BL values, or both represent the same distribution. We define that a p-value greater than 0.95 for any test means acceptance of the hypothesis. The observed first-digit distribution is considered consistent with BL if the hypotheses of two tests are accepted. In addition, the goodness of fit φ introduced by Sambridge et al. (2010) is used to evaluate the difference between the observed distribution and BL:

$$\varphi = \left(1 - \left(\sum_{D=1}^{9} \frac{(f_{obs_d} - f_{BL_d})^2}{f_{BL_d}}\right)^{1/2}\right) \times 100\% \quad (3)$$

where f_{obs} and f_{BL} are the observed digit frequency and theoretical probability of BL, d={1, 2, …, 9}. A value of φ closer to one means that the distribution is closer to the theoretical BL value.

To investigate the occurrence of BL in seismic signals, we examine the relationship between the time series of seismic signals and their corresponding first-digit distribution. This analysis allows us to understand the underlying factors contributing to the emergence of BL during specific processes or periods.

For the processes or periods that follow BL, we examine the relationship between the time series of seismic signals and their first-digit distribution to investigate why BL appears. In the time domain, the raw seismic signal $S(t)$ is a function of time (t) and can be described with an interquartile range iq as magnitude changes of the measurements. Here, the seismic signals before the optimal goodness of fit $\varphi_{optimal}$ were selected to fit an exponential curve for the increased parts (Text S3 for details):

$$S(t) = a * e^{b*t} + c \quad (4)$$

$$iq = Q_{75} - Q_{25} \quad (5)$$

where S is the seismic signals (units are counts), t is time (units are second), and a, b, and c are the coefficients of the exponential function. Q_{75} and Q_{25} are the upper and lower quartile of the data for each window.

Previous studies have demonstrated that datasets with a power law relationship (exponent one) in data pairs satisfy BL, such as the data on many hydrological phenomena (Nigrini & Miller, 2007). In this study, we assume that the seismic data in all one-minute moving windows have this power law distribution, then the data of each window were selected and sorted from smallest to largest (rank order) to calculate α by Equation (6-7) based on its magnitude (Newman, 2005). We subsequently examine whether the seismic data follow the power law with exponent one when BL appears:

$$p(x) = C * x^{-\alpha} \quad (6)$$

$$\alpha = 1 + n\left(\sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}}\right)^{-1} \quad (7)$$

where $p(x)$ is the seismic data in any one-minute moving windows. C and α are the coefficients of the power law function. x_i, and $x_{\min}$ are the i-th data and minimum data in the dataset of length n.

3.3 Data Preparation Debris Flow Detector Implementation

To demonstrate the application of BL, we developed a two-fold debris flow classification detector using seismic data from Illgraben. There are three non-dimensional parameters in our classification model to reduce uncertainty (Figure S2): the ratio between the interquartile range iq at time i and its average value of the previous 20 minutes (R_{iq}), the power exponent at time i (α_i, Equation 6), and the averaged power exponent for d minutes after time i (α_d). We define the debris flow (positive events) by manual interpretation of the seismic data (Text S1 and S4). With a one-minute moving window, all events with all three non-dimensional variables are scanned, and the detector returns either a positive (debris flow) or negative (not debris flow) labels. Finally, in order to test the sensitivity of moving window size or the number of data points for each window, a varying window length from 1s to 600s in a one-second interval was chosen to test the variation of the power law exponent.

Our dataset includes 14 manually marked debris flow events out of a total of debris flow 24 events, we divided the available dataset into a training dataset (24 events, 2013-2014) and a validation dataset (21 events, 2017-2019). The details to define positive and negative cases for training and validation are described in Text S4. We used a confusion matrix to evaluate our detector performance (Beguería, 2006; Staley et al., 2013). The definition of true positive (TP), true negative (TN), false positive (FP), false negative (FN), F1 score (F1), and Threat Score (TS) are given in Text S4. The detector model is considered the best when F1 is one or closest to one. We compare the validation results with an existing random forest model trained with data from 2017 to 2019 recorded by the same seismic network using more than 70 seismic features (Chmiel et al., 2021).

4 Results

4.1 Benford's Law and Seismic Signals

BL was observed in 38 out of the 45 debris flows events, while it was absent for two events from the 2013-2014 GFZ dataset and five events from the 2017-2019 WSL dataset (Figures 2f-2g, S3-S5 and Tables S2-S3). In the debris flow events that follow BL, we observed that both Chi-squared and Kolmogorov-Smirnov tests only accept the hypothesis during the running phase. For example, the debris flow event on 12th July 2014 (Figures 1b and 2a-2d) exhibits $\varphi_{optimal}$ of 87.97% and suggests that the first-digit distribution of seismic signals follows BL. Moreover, the first-digit distributions of station IGB03 and IGB04 are similar to what was observed for this event (Figure 2e). In addition, for other mass movements, the first-digit distribution of the seismic signal generated by the landslide (Figure S8) and the lahar (Figure S9) also exhibit BL. However, our two rockfall cases failed to follow BL (Figures S6-S7).

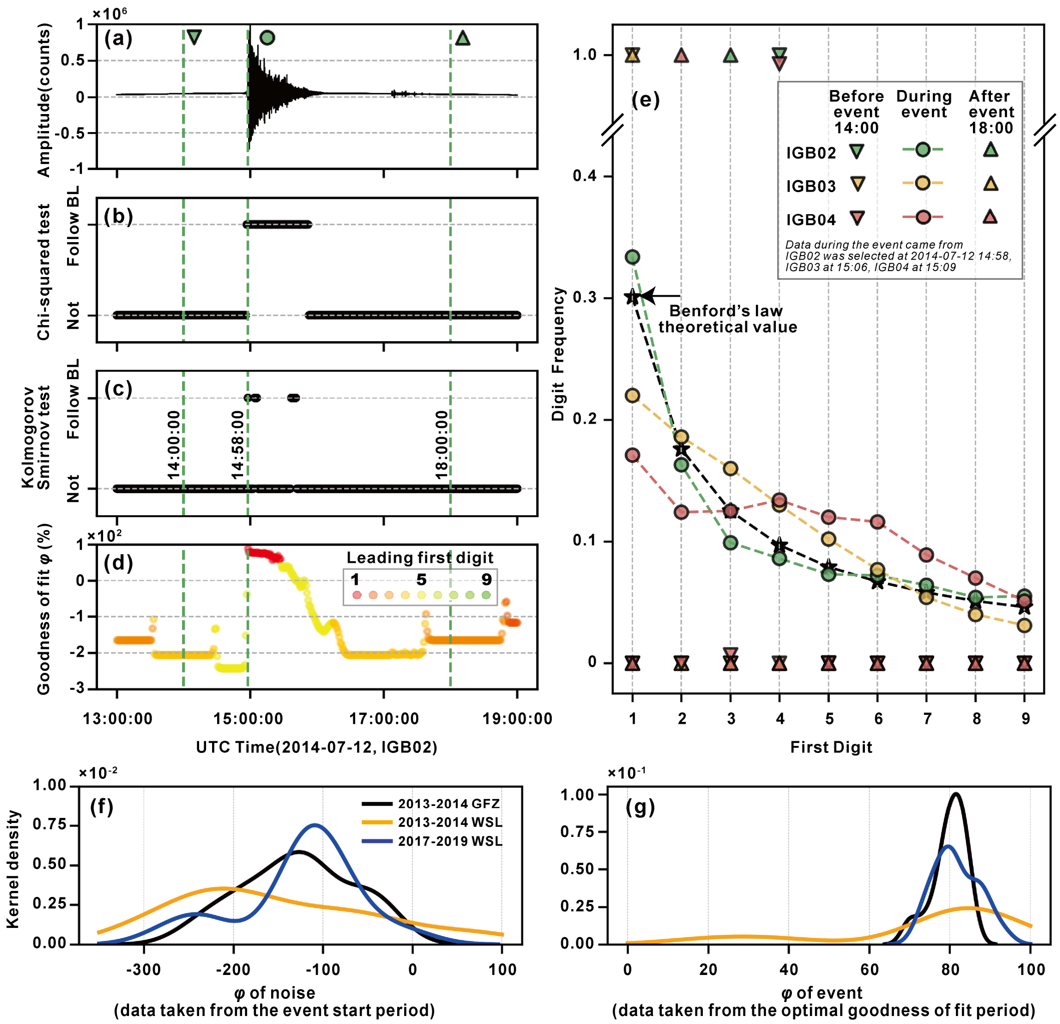


Figure 2 First-digit distribution of seismic signal generated by debris flow. (a) Raw waveform. (b) to (d) are results from the Chi-squared test, Kolmogorov-Smirnov test, goodness of fit. (e) First-digit distribution of different stations during this debris flow. (f) and (g) are kernel density of goodness of fit of the 38 BL-followed events.

For fluvial processes, the 2014 flood case in Illgraben only exhibits a one-minute window obeying BL (Figure S3), but the GLOF obeys BL for a much longer period of 8 minutes (Figure S10). In contrast, the bedload transport cannot be distinguished by BL (Figure S11). Interestingly, the first-digit distribution of seismic waveform generated by long-period seismic signals and when amplitude is close to zero counts fluctuation (named LP0), also follows BL (Figure S12).

4.2 Empirical Analysis for BL

As the debris flow front approaches, both the *iq* and φ rapidly increase in the time series domain (Figure 3a). We observed a good fit between seismic signals S and time t using the exponential function. The kernel density and coefficient of determination R^2 were used to show the exponential fitting difference between the event ($\varphi_{optimal}$ period) and noise (manually labeled event start time period). The averaged R^2 of 2013-2014 WSL and 2017-2019 WSL label events are 0.853 and 0.840, respectively, and the averaged R^2 of 2013-2014 GFZ label events is 0.649, however, the raw amplitude data of the noise period could not be fitted with an averaged R^2 of 0.434 (Figure S13a-b). For instance, event 2014-07-12 has a R^2 of 0.944 (Figure 3a). The same fitting method did not yield an exponential curve for the noise period data of event 2014-07-12 (Figure S14).

In addition, we found that the α for the debris flow, lahar, landslide, and GLOF is 1.10-1.13 (Figure 3b). However, the exponent of two rockfall cases, bedload transport case, and most ambient noise is much higher than one. Values of α close to one could also be observed for ambient noise generated by LP0. The kernel density of α of all 38 BL-obeying debris flow events is much closer to one than the noise period (Figure S13c-d).

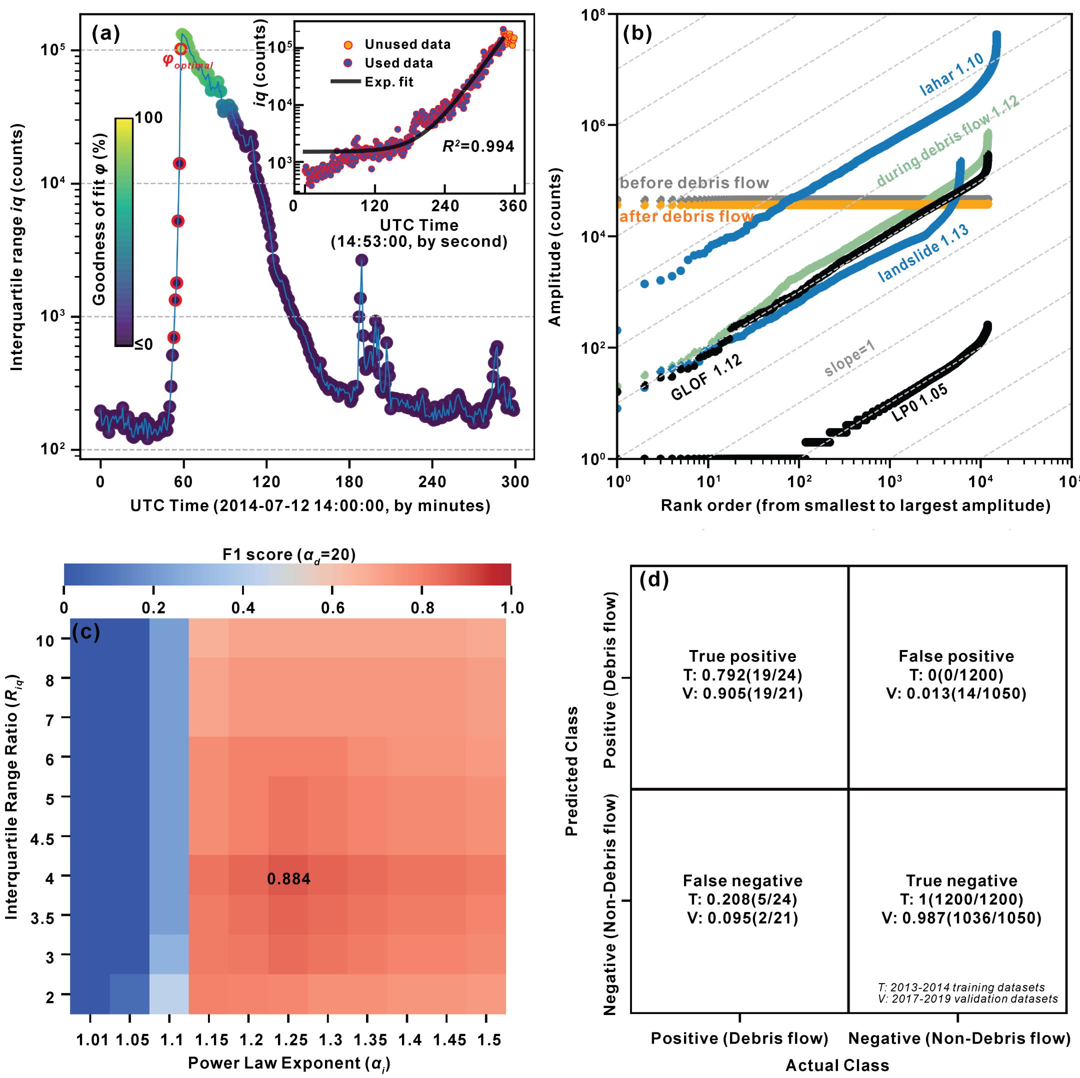


Figure 3 Correlation between seismic signals and BL, and BL-based debris-flow detector performance. (a) *iq* changes in the seismic signals for a debris flow event and exponential fitting (Exp. fitting). (b) Power law relationship in raw amplitude for different events. (c) F1 score for debris-flow event detectors (Figure S15 for details). (d) Confusion matrix of the best detector for training and output of validation dataset under the same parameters.

4.3 Debris-flow Detector Based on BL

The power exponent, which is calculated from the ten-minute average value when the debris flows front approaches station IGB02, converges with an increasing moving window (Figure S16a). During the training procedure, the performance of the detector was examined by the 2013-2014 dataset (Figures 3c and S15-S16). Results show that the power law exponent (α_i=1.25), interquartile range ratio (R_{iq}=4), and event duration (α_d=20 minutes) produce the best detector, yielding an F1 score of 0.884 (TS=0.792, Figure 3c). Under this set of parameters, there are five false-negative events in the training dataset (two events do not follow BL, Figure S17-S21). These three optimal parameters (α_i=1.25, R_{iq}=4, α_d=20) obtained from the training procedure were tested with the validation dataset. We found that the detector produced a higher TPR=0.905 (F1=0.704, TS=0.543) than the training procedure (Figure 3d). However, two false-negative cases for the validation dataset were observed, 14 cases were mislabeled as positive events (false positive) in the validation catalog (Figure 3d).

5 Discussion

5.1 Why Do Some Seismic Datasets Follow BL

The results show that BL is an efficient approach for detecting high-energy mass movements and some fluvial processes with seismic signals. The processes that do follow BL (debris flow, landslide, lahar, and GLOF) usually contain more kinetic energy during the running process than cases that do not follow BL (ambient noise, rockfall, and bedload transport). Interestingly, when the raw waveform is close to zero fluctuation (e.g., one-minute amplitude data between -227 and 342 counts, Figure S12), the observed first-digit distribution could also follow BL. For rockfall and other events that do not follow BL, we argue that the highly attenuated signals or low SNR make it difficult to distinguish between an event and ambient noise in the raw waveform domain. Generally, the low SNR is due to geometric spreading and anelastic attenuation, the energy and amplitude of the signals dissipate during propagation, especially for high-frequency waves (Battaglia, 2003; Tsai & Atiganyanun, 2014).

For seismic datasets, Sambridge et al. (2010) first stated that a sufficient dynamic range may lead the first-digit distribution (e.g., seismic signals generated by an earthquake) to follow BL. However, the claim that regularity and large spread imply BL is not always correct (Berger & Hill, 2011). In theory, datasets crossing several orders of magnitude do not necessarily follow BL (Figure S22a); in practice, we found that the seismic signals that cross two orders of magnitude within one minute follow BL (*iq* 119 counts, Figure S12). For teleseismic events and local seismicity, Díaz et al. (2015) observed using both natural and artificial data that compliance to BL does not depend primarily on the dynamic amplitude range, but rather relates to changes in frequency content. Yet, it is nearly impossible to obtain seismic data in the field that feature only frequency changes without changes in the dynamic range. For seismic signals generated by high-energy mass movements or fluvial processes, signals usually have significant changes in both magnitude (> two orders) and frequency (>1 Hz change in central frequency) when compared to an earthquake (Figures 2a and 3a). In theory, the sole change in the frequency domain does not necessarily cause compliance to BL (Figure S22b). We propose two possible mathematical explanations for the appearance of BL in seismic signals. Firstly, when data adheres to BL, it follows Zipf's law (Newman, 2005). BL appears exactly when the scaling exponent α=1 (Pietronero et al., 2001), so empirical values of the exponent close to one in seismic data will yield compliance to BL. Secondly, BL appears exactly for processes that rise or fall exponentially in time, which mathematically corresponds to a mapping from a linear to a logarithmic space (Cong et al., 2019; Engel & Leuenberger, 2003). Thus, processes that develop with exponential dynamics in time can be expected to follow BL (Figure 3a). This implies that the data from exponentially evolving processes will also follow Zipf's law with a scaling exponent of one.

We suggest that events with an exponentially rising signal can follow BL is primarily caused by their spatial mobility. The amplitude of an approaching seismic source is controlled by the ground quality factor (as an exponential term) and source-receiver distance (as 1 over the square root of the distance) (Burtin et al., 2016). As long as the distance at which a process emits sufficient energy to be detected by a seismometer is much larger than the channel-sensor distance (Figure 1a), a fast-moving mass will produce a signal sufficiently close to exponential increase (Dietze et al., 2022) to be in agreement with BL. In other words, BL is an efficient detector of fast approaching seismic sources at the landscape scale.

5.2 Application of BL as early warning tool

BL is a computationally cheap and novel approach to detect debris flow and establish real-time warning systems. Since only the raw data need to be counted, the computation time for parameter preparation and model evolution is strongly reduced, e.g., our validation process could be completed in 113 seconds (Figure S2). We expect that BL can be applied to different sites without a change in parameter values, because of the predefined non-dimensional variables and their general applicability. Therefore, we suggest that the dimensionless detector input parameters are independent of catchment geometry and seismic station characteristics. Furthermore, our approach could be a simplified version of an early warning system for triggering or turning on high-power supply and data transmission devices to catch events, such as radar and laser, for full-scale warning. In practice, a BL-based early warning system can be implemented using data from two or three seismic stations along the main flow path to detect and cross-validate events.

Our purpose in this paper is to explore the potential of BL as a prerequisite to developing an operational event detector or warning system for debris flows, which can be adapted to other processes as well. An efficient real-time warning system requires the rapid detection of the event of interest, and signal processing plays a critical role in validating seismology-driven warning systems (Arattano et al., 2014, 2016; Coviello et al., 2015). By using BL, high-energy processes lasting longer than a few minutes can be reliably distinguished from background noise. Without changing input parameters, our detector achieves a detection accuracy of 0.905 and 0.982 for both debris flows and non-debris-flow events in the validation catalog (Figure 3d). This is similar to the detection accuracy from a random forest model calibrated in the same catchment (Chmiel et al., 2021), which gives 0.83 and 0.94 for debris flows before or after check dam 1 (marked in Figure 1a) and 0.92 for non-debris-flow events, respectively. Any supervised machine-learning-based model, including those based on random forests, requires a large training dataset from multiple seismic stations and many seismic features. The efficacy of our debris-flow detector is at least comparable to the random forest model, but does not require recalibration of parameters. Furthermore, the false-positive example of our model in the validation catalog can be filtered out using data from multiple seismic stations (Figure S23), and a full seismic network can improve true positive detection accuracy (Figures S24-S25).

More mass movements are needed to help understanding the scope of application of BL and the detector proposed in this paper needs to be further explored. This study suggests that a single seismic station could efficiently detect events such as debris flow that move continuously in the channel. However, the BL approach neglects frequency domain information, which could be used to improve the identification of other high-energy mass movements or fluvial processes type.

6 Conclusion

Detecting events of interest from seismic signals to establish early warning systems is critical for hazard mitigation. In this study, we demonstrated that the first-digit distribution of seismic signals generated by some high-energy mass movements and fluvial processes follows Benford's law. Our detector model offers a less computationally intensive and novel approach for extracting anomalous energetic events, such as debris flows and landslides, from massive seismic signals. Moreover, the high-energy mass movement detector provides a promising strategy for building warning systems using seismic signals to mitigate hazards. In the future, we will collect more mass movements to calculate their first-digit distribution and develop a seismic network-based detector system and implement our method for real-time detection.

Acknowledgments

The authors thank Dr. Lucia Capra provide the lahar data. This study is funded by the State Key Laboratory of Geohazard Prevention and Geoenvironment Protection Open Fund (SKLGP2023K003) and National Natural Science Foundation of China (42120104002).

Open Research

The seismic data between 2013 and 2014 is available at https://doi.org/10.14470/4W615776. Please refer to *Chmiel, M., Walter, F., Wenner, M., Zhang, Z., McArdell, B. W., & Hibert, C. (2021). Machine learning improves debris flow warning. Geophysical Research Letters, 48, e2020GL090874. https://doi.org/10.1029/2020GL090874* for 2017 to 2019 seismic data. Our figures were generated using R ESEIS, Obspy, Matplotlib and Seaborn. All seismic data processing codes are available in *https://github.com/Nedasd/Benfords-law-in-environmental-seismology.git*

References

Arattano, M., Abancó, C., Coviello, V., & Hürlimann, M. (2014). Processing the ground vibration signal produced by debris fl ows : the methods of amplitude and impulses compared. *Computers and Geosciences*, *73*, 17–27. https://doi.org/10.1016/j.cageo.2014.08.005

Arattano, M., Coviello, V., Abancó, C., Hürlimann, M., & Mcardell, B. W. (2016). Methods of Data Processing for Debris Flow Seismic Warning. *International Journal of Erosion Control Engineering*, *9*(3).

Badoux, A., Graf, C., Rhyner, J., Kuntner, R., & McArdell, B. W. (2009). A debris-flow alarm system for the Alpine Illgraben catchment: Design and performance. *Natural Hazards*, *49*(3), 517–539. https://doi.org/10.1007/s11069-008-9303-x

Battaglia, J. (2003). Location of seismic events and eruptive fissures on the Piton de la Fournaise volcano using seismic amplitudes. *Journal of Geophysical Research*, *108*(B8), 1–14. https://doi.org/10.1029/2002jb002193

Baum, R. L., & Godt, J. W. (2010). Early warning of rainfall-induced shallow landslides and debris flows in the USA. *Landslides*, *7*(3), 259–272. https://doi.org/10.1007/s10346-009-0177-0

Beguería, S. (2006). Validation and evaluation of predictive models in hazard assessment and risk management. *Natural Hazards*, *37*(3), 315–329. https://doi.org/10.1007/s11069-005-5182-6

Benford, F. (1938). The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society*. Retrieved from http://www.jstor.org/stable/984802

Berger, A., & Hill, T. P. (2011). Benford's law strikes back: No simple explanation in sight for mathematical gem. *The Mathematical Intelligencer*, *33*(1), 85.

Le Breton, M., Bontemps, N., Guillemot, A., Baillet, L., & Larose, É. (2021). Landslide monitoring using seismic ambient noise correlation: challenges and applications. *Earth-Science Reviews*, *216*(December 2020). https://doi.org/10.1016/j.earscirev.2021.103518

Burtin, A., Hovius, N., McArdell, B. W., Turowski, J. M., & Vergne, J. (2014). Seismic constraints on dynamic links between geomorphic processes and routing of sediment in a steep mountain catchment. *Earth Surface Dynamics*, *2*(1), 21–33. https://doi.org/10.5194/esurf-2-21-2014

Burtin, A., Hovius, N., & Turowski, J. M. (2016). Seismic monitoring of torrential and fluvial processes. *Earth Surface Dynamics*, *4*(2), 285–307. https://doi.org/10.5194/esurf-4-285-2016

Capra, L., Coviello, V., Borselli, L., Márquez-Ramírez, V. H., & Arámbula-Mendoza, R. (2018). Hydrological control of large hurricane-induced lahars: Evidence from rainfall-runoff modeling, seismic and video monitoring. *Natural Hazards and Earth System Sciences*, *18*(3), 781–794. https://doi.org/10.5194/nhess-18-781-2018

Castañeda, G. (2011). La ley de Benford y su aplicabilidad en el análisis forense de resultados electorales. *Politica y Gobierno*, *18*(2), 297–330.

Chmiel, M., Walter, F., Wenner, M., Zhang, Z., McArdell, B. W., & Hibert, C. (2021). Machine Learning Improves Debris Flow Warning. *Geophysical Research Letters*, *48*(3), 1–11. https://doi.org/10.1029/2020GL090874

Cho, W. K. T., & Gaines, B. J. (2007). Breaking the (Benford) law: Statistical fraud detection in campaign finance. *American Statistician*, *61*(3), 218–223. https://doi.org/10.1198/000313007X223496

Comiti, F., Marchi, L., Macconi, P., Arattano, M., Bertoldi, G., Borga, M., et al. (2014). A new monitoring station for debris flows in the European Alps: First observations in the Gadria basin. *Natural Hazards*, *73*(3), 1175–1198. https://doi.org/10.1007/s11069-014-1088-5

Cong, M., Li, C., & Ma, B. Q. (2019). First digit law from Laplace transform. *Physics Letters, Section A: General, Atomic and Solid State Physics*, *383*(16), 1836–1844. https://doi.org/10.1016/j.physleta.2019.03.017

Cook, K. L., & Dietze, M. (2022). Seismic Advances in Process Geomorphology. *Annual Review of Earth and Planetary Sciences*, *50*(1), 183–204. https://doi.org/10.1146/annurev-earth-032320-085133

Cook, K. L., Andermann, C., Gimbert, F., Adhikari, B. R., & Hovius, N. (2018). Glacial lake outburst floods as drivers of fluvial erosion in the Himalaya. *Science*, *362*(6410), 53–57. https://doi.org/10.1126/science.aat4981

Cook, K. L., Rekapalli, R., Dietze, M., Pilz, M., Cesca, S., Rao, N. P., et al. (2021). Detection and potential early warning of catastrophic flow events with regional seismic networks. *Science*, *374*(6563), 87–92. https://doi.org/10.1126/science.abj1227

Coussot, P., & Meunier, M. (1996). Recognition, classification and mechanical description of debris flows. *Earth-Science Reviews*, *40*(3–4), 209–227. https://doi.org/10.1016/0012-8252(95)00065-8

Coviello, V., Arattano, M., & Turconi, L. (2015). Detecting torrential processes from a distance with a seismic monitoring network. *Natural Hazards*, *78*(3), 2055–2080. https://doi.org/10.1007/s11069-015-1819-2

Coviello, V., Arattano, M., Comiti, F., Macconi, P., & Marchi, L. (2019). Seismic Characterization of Debris Flows: Insights into Energy Radiation and Implications for Warning. *Journal of Geophysical Research: Earth Surface*, *124*(6), 1440–1463. https://doi.org/10.1029/2018JF004683

Dai, F. C., Lee, C. F., & Ngai, Y. Y. (2002). Landslide risk assessment and management: An overview. *Engineering Geology*, *64*(1), 65–87. https://doi.org/10.1016/S0013-7952(01)00093-X

Dammeier, F., Moore, J. R., Hammer, C., Haslinger, F., & Loew, S. (2016). Automatic detection of alpine rockslides in continuous seismic data using hidden Markov models. *Journal of Geophysical Research: Earth Surface*, *121*(2), 351–371. https://doi.org/10.1002/2015JF003647

Díaz, J., Gallart, J., & Ruiz, M. (2015). On the ability of the Bedford's law to detect earthquakes and discriminate seismic signals. *Seismological Research Letters*, *86*(1), 192–201. https://doi.org/10.1785/0220140131

Dietze, M., Mohadjer, S., Turowski, J. M., Ehlers, T. A., & Hovius, N. (2017). Seismic monitoring of small alpine rockfalls-validity, precision and limitations. *Earth Surface Dynamics*, *5*(4), 653–668. https://doi.org/10.5194/esurf-5-653-2017

Dietze, M., Hoffmann, T., Bell, R., Schrott, L., & Hovius, N. (2022). A Seismic Approach to Flood Detection and Characterization in Upland Catchments Geophysical Research Letters. *Geophysical Research Letters*, *49*, 1–8. https://doi.org/10.1029/2022GL100170

Dowling, C. A., & Santi, P. M. (2014). Debris flows and their toll on human life: A global analysis of debris-flow fatalities from 1950 to 2011. *Natural Hazards*, *71*(1), 203–227. https://doi.org/10.1007/s11069-013-0907-4

Ekström, G., & Stark, C. P. (2013). Simple scaling of catastrophic landslide dynamics. *Science*, *339*(6126), 1416–1419. https://doi.org/10.1126/science.1232887

Engel, H. A., & Leuenberger, C. (2003). Benford's law for exponential random variables. *Statistics and Probability Letters*, *63*(4), 361–365. https://doi.org/10.1016/S0167-7152(03)00101-9

Fan, X., Scaringi, G., Korup, O., West, A. J., van Westen, C. J., Tanyas, H., et al. (2019). Earthquake-Induced Chains of Geologic Hazards: Patterns, Mechanisms, and Impacts. *Reviews of Geophysics*, *57*(2), 421–503. https://doi.org/10.1029/2018RG000626

Farin, M., Tsai, V. C., Lamb, M. P., & Allstadt, K. E. (2019). A physical model of the high-frequency seismic signal generated by debris flows. *Earth Surface Processes and Landforms*, *44*(13), 2529–2543. https://doi.org/10.1002/esp.4677

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, *18*(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Fuchs, S., Heiss, K., & Hübl, J. (2007). Natural Hazards and Earth System Sciences Towards an empirical vulnerability function for use in debris flow risk assessment. *Hazards Earth Syst. Sci*, *7*(1), 495–506. Retrieved from www.nat-hazards-earth-syst-sci.net/7/495/2007/

Geyer, A., & Martí, J. (2012). Applying Benford's law to volcanology. *Geology*, *40*(4), 327–330. https://doi.org/10.1130/G32787.1

Govi, M., Maraga, F., & Moia, F. (1993). Seismic detectors for continuous bed load monitoring in a gravel stream. *Hydrological Sciences Journal*, *38*(2), 123–132. https://doi.org/10.1080/02626669309492650

Gregoretti, C., Degetto, M., Bernard, M., Crucil, G., Pimazzoni, A., De Vido, G., et al. (2016). Runoff of small rocky headwater catchments: Field observations and hydrological modeling. *Water Resources Research*, *52*(10), 8138–8158. https://doi.org/10.1111/j.1752-1688.1969.tb04897.x

Guzzetti, F., Gariano, S. L., Peruccacci, S., Brunetti, M. T., Marchesini, I., Rossi, M., & Melillo, M. (2020). Geographical landslide early warning systems. *Earth-Science Reviews*, *200*(September 2019), 102973. https://doi.org/10.1016/j.earscirev.2019.102973

Hammer, C., Beyreuther, M., & Ohrnberger, M. (2012). A seismic-event spotting system for volcano fast-response systems. *Bulletin of the Seismological Society of America*, *102*(3), 948–960. https://doi.org/10.1785/0120110167

Hibert, C., Michéa, D., Provost, F., Malet, J. P., & Geertsema, M. (2019). Exploration of continuous seismic recordings with a machine learning approach to document 20 yr of landslide activity in Alaska. *Geophysical Journal International*, *219*(2), 1138–1147. https://doi.org/10.1093/gji/ggz354

Holub, M., & Hübl, J. (2008). Local protection against mountain hazards - State of the art and future needs. *Natural Hazards and Earth System Science*, *8*(1), 81–99. https://doi.org/10.5194/nhess-8-81-2008

Huebl, J., & Fiebiger, G. (2015). *Debris-flow mitigation measures*. Debris-flow hazards and related phenomena. https://doi.org/10.1007/3-540-27129-5

Hürlimann, M., Coviello, V., Bel, C., Guo, X., Berti, M., Graf, C., et al. (2019). Debris-flow monitoring and warning: Review and examples. *Earth-Science Reviews*, *199*(October), 102981. https://doi.org/10.1016/j.earscirev.2019.102981

Joannes-Boyau, R., Bodin, T., Scheffers, A., Sambridge, M., & May, S. M. (2015). Using Benford's law to investigate natural hazard dataset homogeneity. *Scientific Reports*, *5*, 1–8. https://doi.org/10.1038/srep12046

Kean, J. W., Staley, D. M., Lancaster, J. T., Rengers, F. K., Swanson, B. J., Coe, J. A., et al. (2019). Inundation, flow dynamics, and damage in the 9 January 2018 Montecito debris-flow event, California, USA: Opportunities and challenges for post-wildfire risk assessment. *Geosphere*, *15*(4), 1140–1163. https://doi.org/10.1130/GES02048.1

Ley, E. (1996). On the Peculiar Distribution of the U.S. Stock Indexes' Digits. *American Statistician*. https://doi.org/10.1080/00031305.1996.10473558

Marra, F., Nikolopoulos, E. I., Creutin, J. D., & Borga, M. (2016). Space–time organization of debris flows-triggering rainfall and its effect on the identification of the rainfall threshold relationship. *Journal of Hydrology*, *541*, 246–255. https://doi.org/10.1016/j.jhydrol.2015.10.010

McArdell, B. W., Bartelt, P., & Kowalski, J. (2007). Field observations of basal forces and fluid pore pressure in a debris flow. *Geophysical Research Letters*, *34*(7), 2–5. https://doi.org/10.1029/2006GL029183
Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J. C. J. H., Bates, P., et al. (2021). Causes, impacts and patterns of disastrous river floods. *Nature Reviews Earth and Environment*, *2*(9), 592–609. https://doi.org/10.1038/s43017-021-00195-3
Nemčok, A., Pašek, J., & Rybář, J. (1972). Classification of landslides and other mass movements. *Rock Mechanics Felsmechanik Mécanique Des Roches*, *4*(2), 71–78. https://doi.org/10.1007/BF01239137
Newcomb, S. (1881). Note on the Frequency of Use of the Different Digits in Natural Numbers. *American Journal of Mathematics*, *4*(1/4), 39. https://doi.org/10.2307/2369148
Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*(5), 323–351. https://doi.org/10.1080/00107510500052444
Nigrini, M. J., & Miller, S. J. (2007). Benford's Law applied to hydrology data - Results and relevance to other geophysical data. *Mathematical Geology*, *39*(5), 469–490. https://doi.org/10.1007/s11004-007-9109-5
Nikolopoulos, E. I., Borga, M., Creutin, J. D., & Marra, F. (2015). Estimation of debris flow triggering rainfall: Influence of rain gauge density and interpolation methods. *Geomorphology*, *243*, 40–50. https://doi.org/10.1016/j.geomorph.2015.04.028
Pietronero, L., Tosatti, E., Tosatti, V., & Vespignani, A. (2001). Explaining the uneven distribution of numbers in nature: The laws of Benford and Zipf. *Physica A: Statistical Mechanics and Its Applications*, *293*(1–2), 297–304. https://doi.org/10.1016/S0378-4371(00)00633-6
Provost, F., Hibert, C., & Malet, J. P. (2017). Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier. *Geophysical Research Letters*, *44*(1), 113–120. https://doi.org/10.1002/2016GL070709
Regmi, N. R., Giardino, J. R., McDonald, E. V., & Vitek, J. D. (2015). A Review of Mass Movement Processes and Risk in the Critical Zone of Earth. *Developments in Earth Surface Processes*, *19*, 319–362. https://doi.org/10.1016/B978-0-444-63369-9.00011-2
Sambridge, M., Tkalčić, H., & Jackson, A. (2010). Benford's law in the natural sciences. *Geophysical Research Letters*, *37*(22), 1–5. https://doi.org/10.1029/2010GL044830
Schimmel, A., & Hübl, J. (2016). Automatic detection of debris flows and debris floods based on a combination of infrasound and seismic signals. *Landslides*, *13*(5), 1181–1196. https://doi.org/10.1007/s10346-015-0640-z
Schimmel, A., Kogelnig, A., Vilajosana, I., & Mcardell, B. W. (2013). A review on acoustic monitoring of debris flow. *International Journal of Safety and Security Engineering*, *3*(2), 105–115. https://doi.org/10.2495/SAFE-V3-N2-105-115
Schöpa, A., Chao, W. A., Lipovsky, B. P., Hovius, N., White, R. S., Green, R. G., & Turowski, J. M. (2018). Dynamics of the Askja caldera July 2014 landslide, Iceland, from seismic signal analysis: Precursor, motion and aftermath. *Earth Surface Dynamics*, *6*(2), 467–485. https://doi.org/10.5194/esurf-6-467-2018
Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M., & Laber, J. L. (2013). Objective definition of rainfall intensity-duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides*, *10*(5), 547–562. https://doi.org/10.1007/s10346-012-0341-9

Sun, W., & Tkalčić, H. (2022). Repetitive marsquakes in Martian upper mantle. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-29329-x

Tiwari, A., Sain, K., Kumar, A., Tiwari, J., Paul, A., Kumar, N., et al. (2022). Potential seismic precursors and surficial dynamics of a deadly Himalayan disaster : an early warning approach. *Scientific Reports*, 1–13. https://doi.org/10.1038/s41598-022-07491-y

Tsai, V. C., & Atiganyanun, S. (2014). Green's functions for surface waves in a generic velocity structure. *Bulletin of the Seismological Society of America*, *104*(5), 2573–2578. https://doi.org/10.1785/0120140121

Wei, S. C., & Liu, K. F. (2020). Automatic debris flow detection using geophones. *Landslides*, *17*(2), 349–359. https://doi.org/10.1007/s10346-019-01258-9

Wilson, R. C., & Wieczorek, G. F. (1995). Rainfall Thresholds for the Initiation of Debris Flows at La Honda, California. *Environmental & Engineering Geoscience*, *1*(1), 11–27. Retrieved from https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civil wars_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625

Supporting Information for

Benford's law as mass movement detector in seismic signals

Qi Zhou[1, 2†], Hui Tang[1], Jens M. Turowski[1], Jean Braun[1, 2], Michael Dietze[3,1], Fabian Walter[4], Ci-Jian Yang[1, 5], Sophie Lagarde[1, 2]

[1]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam,

[2]Institute of Geosciences, University of Potsdam, Potsdam,

[3]Faculty of Geoscience and Geography, Georg-August-Universität Göttingen, Göttingen,

[4]Swiss Federal Institute for Forest, Snow and Landscape Research, Zürich,

[5]Department of Geography, National Taiwan University, Taipei,

Corresponding author: Qi Zhou (qi.zhou@gfz-potsdam.de, F427, Telegrafenberg, 14473 Potsdam, Germany)

Contents of this file

Introduction

This file includes supplementary text, figures, tables, and code for the manuscript titled *Benford's law as mass movement detector in seismic signals*.

Supplemental Text1 presents a comprehensive description of the methods employed to label debris flow events between the years 2013 and 2014. **Supplemental Text 2** illustrates the seismic data source and the catalog of debris flow events from 2017 to 2019. **Supplemental Text3** provides an elaborate explanation of the methods employed to fit the exponential curve as debris flow approaches the seismic station IGB02. Lastly, **Supplemental Text4** offers additional information regarding the definition of the training and validation classes, as well as details about the detector model elevation.

Figure S1 demonstrates the effects of different seismic data processing methods on Benford's Law (BL). **Figure S2** shows the input parameters of the debris flow detector model. **Figures S3** to **S4** and **Figure S5** display the BL results of training events between 2013 and 2014 and between 2017 and 2019, respectively, with a focus on partial event representations. **Figures S6** to **S9** present the BL results of different mass movements, and **Figures S10** to **S12** exhibit the BL results of various fluvial processes. **Figure S13** show the kernel density for exponential fitting and power law exponent of noise and 38 BL-followed events. **Figure S14** display an example of exponential fitting curve results. **Figure S15** demonstrates the sensitivity of the debris flow detector model to different input parameters. **Figure S16** show the variation of power law exponent with different moving window sizes and detectors receiver operating characteristic. **Figures S17** to **S21** depict false-negative events encountered during the training process. **Figure S22** illustrates the relationship between BL and dataset range/signal frequency. **Figures S23** to **S25** highlight the potential of using a seismic array to improve detection performance, reduce the accuracy of false negative and improve the accuracy of true positive.

Table S1 provides parameters of seismometer stations for 2013-2014 and 2017-2019. **Table S2** to **S3** offer detailed information on start and end times, and whether BL was followed for the 24 training and 21 validation events. **Table S4** presents the seismic data source for different mass movements and fluvial processes to examine compliance with BL. **Table S5** displays the exponential fitting coefficients for all 45 debris flow events.

Text S1.

Methods to label debris flow training events (2013-2014)

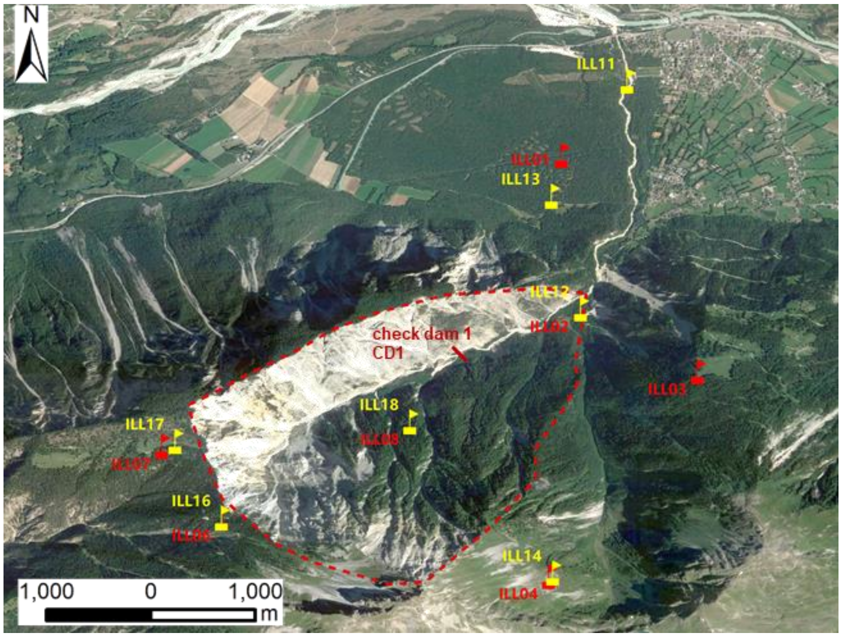
In order to build a training events catalog using data from 2013 to 2014, we utilized the data recorded by the seismic station IGB02 (located at coordinates 46.2863735252704, 7.62780978682399, elevation 933m, and 60m away from the trunk channel) to extract debris-flow events. The Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) warning systems provides traces for labelling ten debris flow events that occurred during the year 2013-2014, and we manually selected the event start and end time based on the seismic signals of station IGB02 (Table S2).

For debris flow events that may have occurred but were not recorded by the warning systems, we used the data from the seismic station IGB02 to extract events. The data were filtered between 1 and 90 Hz, and the spectrograms were plotted on a daily basis. Then the spectrograms were manually analyzed to retrieve the events. Here we refer to the methodology described in Burtin et al. (2014, 2016), Chmiel et al. (2021), Belli et al., (2022) and mainly focus on two main features that (1) the event duration of waveforms (debris flow should last more than 20 minutes), and (2) frequency features of the spectrogram (the characteristic frequency between 1 and 50 Hz, and peak frequency is around 7 Hz). For debris flows, we do not qualify an event corresponding to the surge, one debris flow could have different surges, and all these surges were merged as one event. Please refer to Table S2 for detailed catalog information. Because we do not have other data to validate the manually extracted events, it may introduce uncertainty.

Text S2.

Data source of debris flow validation events (2017-2019)

Between 2017 and 2019, there were 22 events collected by WSL (Chmiel et al., 2021), while the event 2017-05-19 is not available for ILL02. This event does not be included in the validation dataset. For event details, please refer to the supporting information for "*Machine Learning improves debris flow warning*". https://doi.org/10.1029/2020GL090874



Locations of seismometers at the Illgraben catchment from 2017 (red), 2018-2019 (yellow).

Text S3.

Methods to fit the exponential curve

For each BL-followed debris flow event, we first identify the one-minute time window t_i corresponding to the optimal goodness of fit within the manually labeled start and end times. Since the speed of debris flows can vary, the time required to record their approach by the seismometer may differ. To account for this variability, we select different lengths of raw waveforms (1 to 5 minutes before t_i, by minute) to fit an exponential curve. We then calculate the interquartile range (*iq*) of the raw waveforms using a 1-second sub-window. Next, we extract the dataset $\boldsymbol{X}$ during and before t_i. In cases where *iq* decreases at some points, we feed the data before the optimal of the $\boldsymbol{X}$ dataset to the *scipy.optimize.curve_fit* algorithm to obtain the exponential curve (Equation 1) and correlation coefficient R^2. Finally, we select the optimal fitting curve (optimal R^2) based on all fitting, and the results are listed in Table S5.

$$S(t) = a * e^{b*t} + c \tag{1}$$

where S is the seismic signals (unit by counts), t is time (unit by second), and a, b, and c are the coefficient of the exponential function. If b is smaller than zero, then R^2 is zero.

The codes are available in content **Code S1**.

Text S4.

Define training and validation class

Our dataset has two labels: positive (debris flow) and negative (not debris flow). During the training process, the 24 events (Tables S2) between 2013 and 2014 were labeled positive based on data from station IGB02. An additional 1200 negative cases were randomly selected from outside the event at random start times and durations (between

20 minutes and 6 hours). The ratio between the positive and negative cases (P2N) is 0.02. During the validation process, the 21 debris-flow events between 2017 and 2019 (Tables S3) were labeled positive based on stations ILL02 and ILL12. A further 1050 negative cases were randomly selected using the same method as the training procedure with the same P2N ratio. The validation dataset of 1071 events was processed using the optimal detector determined during training, and the validation results with an existing random forest model trained with data from 2017 to 2019 recorded by the same seismic network using more than 70 seismic features (Chmiel et al., 2021).

Define TP, TN, FP, and FN class

We define the confusion matrix as follows: A true positive (TP) is an event that is classified as a debris flow by observation and our detector. A true negative (TN) is an event that is not classified as a debris flow by either observation or our detector. An event is labeled false positive (FP) when it is labeled as a debris flow by our detector but considered as a non-debris-flow event in observation. An event is labeled as false negative (FN) when an observed debris flow is not classified as debris flow by the detector.

Evaluate detector model

A good classification model should maximize the number of true-positive predictions and minimize the number of false-positive and false-negative predictions. The overall performance of the classifier is quantified as a true-positive rate (*TPR*, recall rate or sensitivity), false-positive rate (*FPR*, fall-out rate), false-negative rate (*FNR*, miss rate), and true-negative rate (*TNR*, specificity) which are calculated as Equation 2-5:

$$TPR = \frac{TP}{TP + FN} \quad (2)$$

$$FPR = \frac{FP}{FP + TN} \quad (3)$$

$$TNR = \frac{TN}{TN + FP} \quad (4)$$

$$FNR = \frac{FN}{FN + TP} \quad (5)$$

where TP, TN, FP, and FN are the number of true-positive, true-negative, false-positive and false-negative events.

We calculated the F-score or F1 score (F1) and the Threat Score (TS), which are commonly used to represent the predictive capability of classification models (Equation 6-7). The F1 measures the accuracy of a binary classification model and is a harmonic mean of precision and recall. When F1 gets the highest possible value (one), it means that the precision and recall are perfect, and F1 receives the lowest value (zero) when either the precision or the recall is zero. The TS is a measure of the overall performance of the classification model. In a perfect model, the threat value would be equal to one, with each false prediction (false-negative or false-positive event) decreasing the value of the threat score.

$$F1 = \frac{2TP}{2TP + FN + FP} \tag{6}$$

$$TS = \frac{TP}{TP + FN + FP} \tag{7}$$

The results of F1 and TS are displayed in Figure S15, the detector model is considered the best when F1 is maximum.

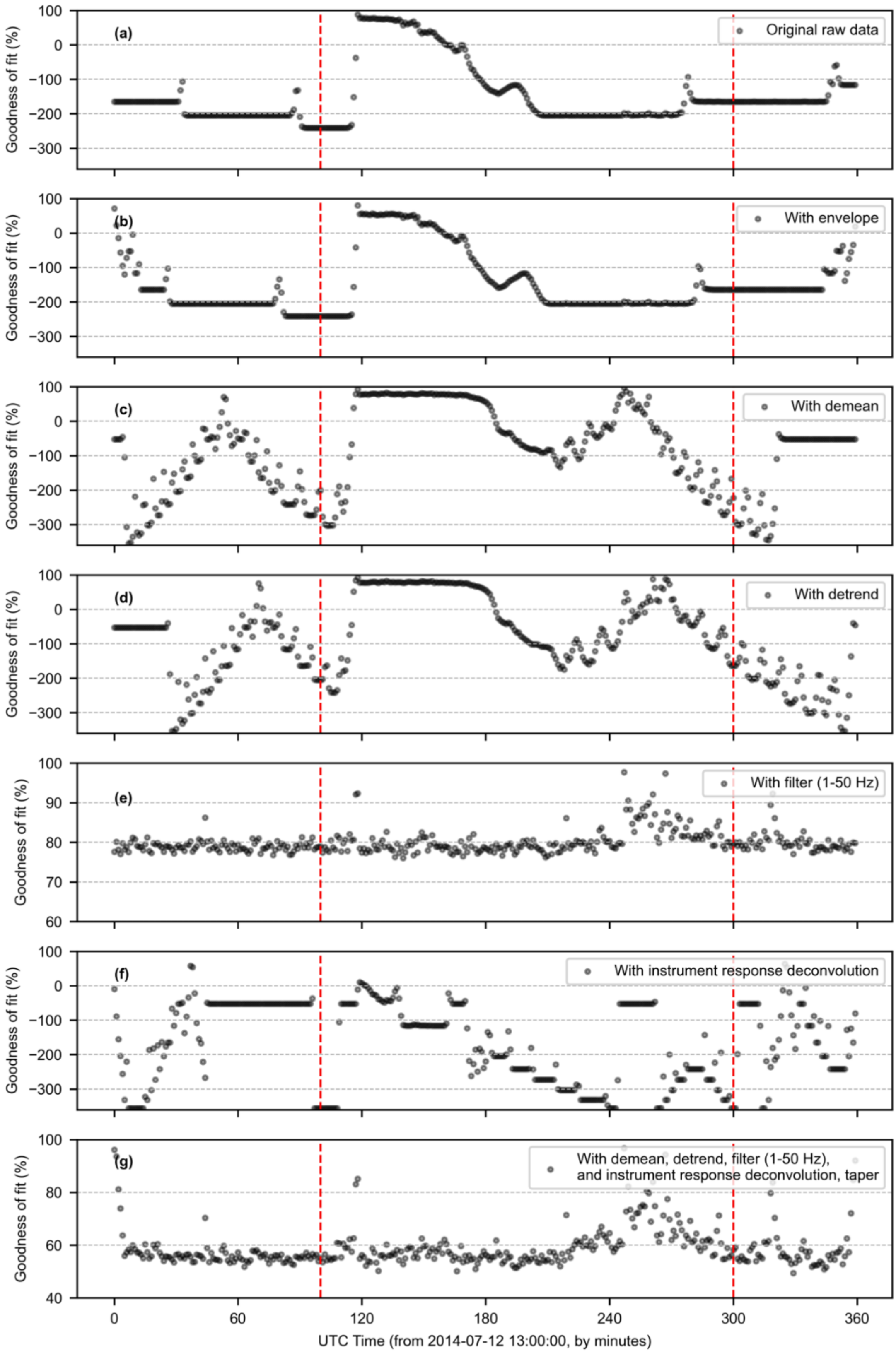


Figure S1. Effects of data processing methods on BL, coding by R ESEIS Dietze (2018) . The red dashed vertical lines are manually labeled start and end time. The codes are available in **4. Codes**.

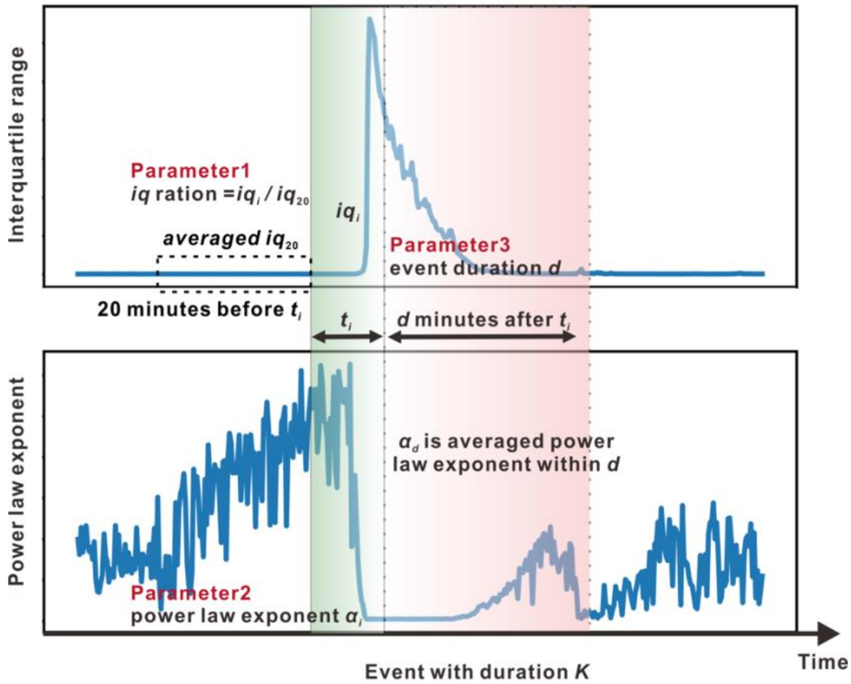


Figure S2. The input dimensionless parameters and range of event detector**.**

The input parameters of detector in Figure S2 are listed below:

$$R_{iq} = iq_i / iq_{i20}, \tag{8}$$

where iq_i is the parameter 1, interquartile range ratio of time i and iq_{i20} is the average interquartile range of 20 minutes before the time i.

Range of input parameters		
Interquartile range (iq) ratio	Power law exponent α_i	Event duration d (unit: minute) or α_d
From 2 to 10	From 1.01 to 1.50	From 3 to 20
Optimal parameters from training data		
iq = 4	α_i = 1.25	d = 20

The validation process with optimal parameters could be completed in 113 seconds (do not count parameters calculating time, and codes were operated with 1 node and 48 G memory) via GFZ Cluster.

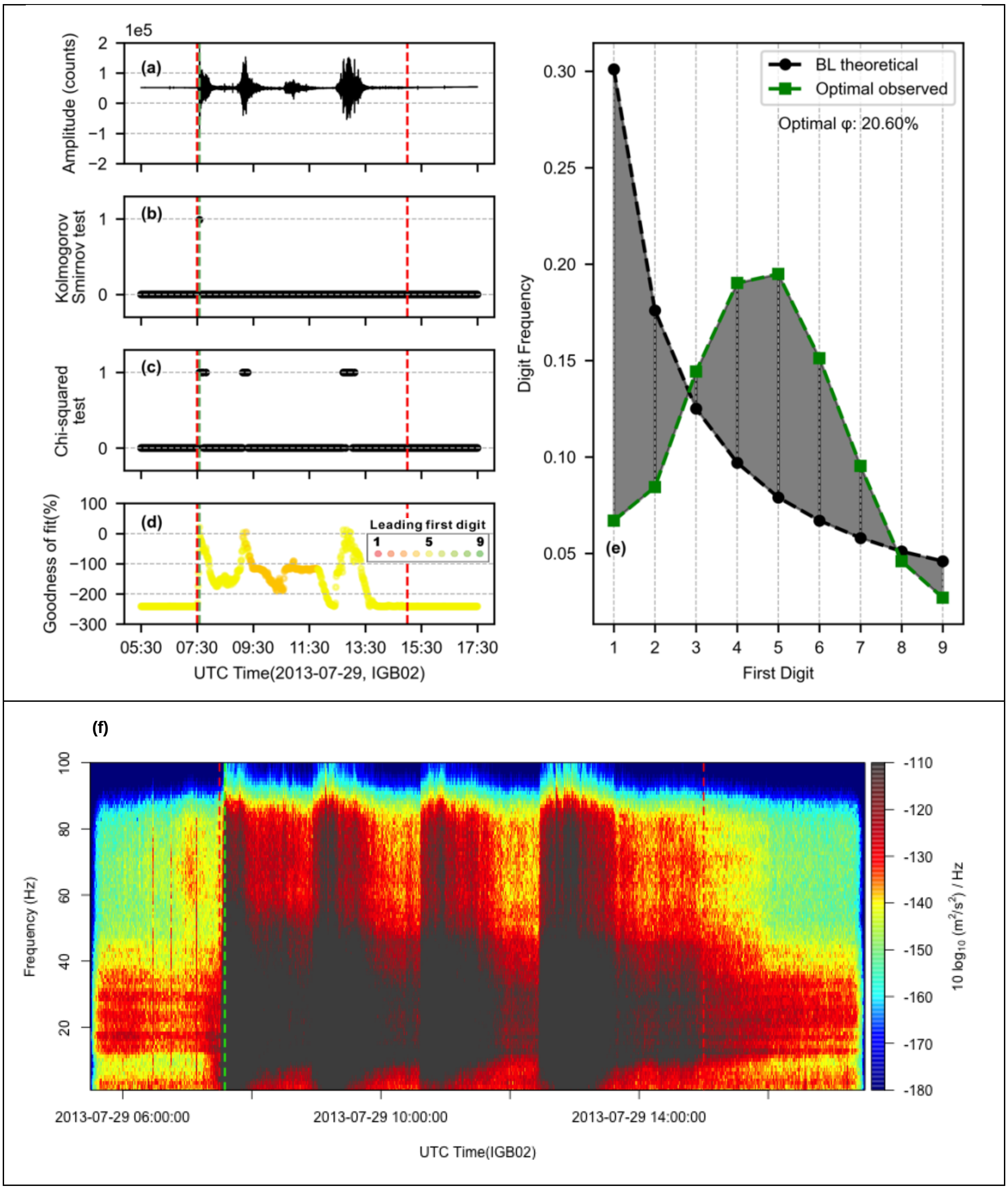


Figure S3. Raw waveform (a) and spectrogram (f) generated by a **flood** (2013-07-29, training events, WSL label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

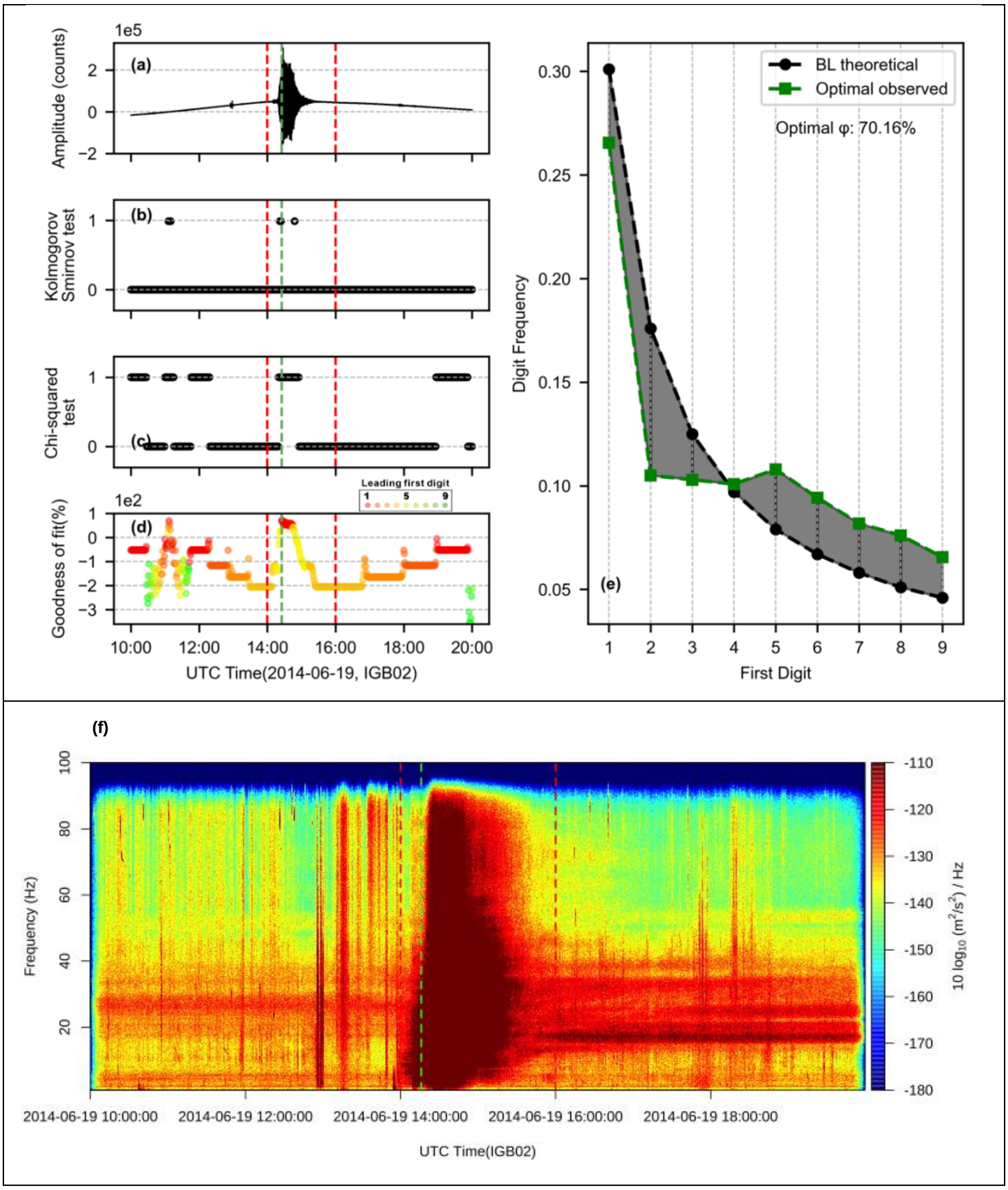


Figure S4. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (2014-06-19, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

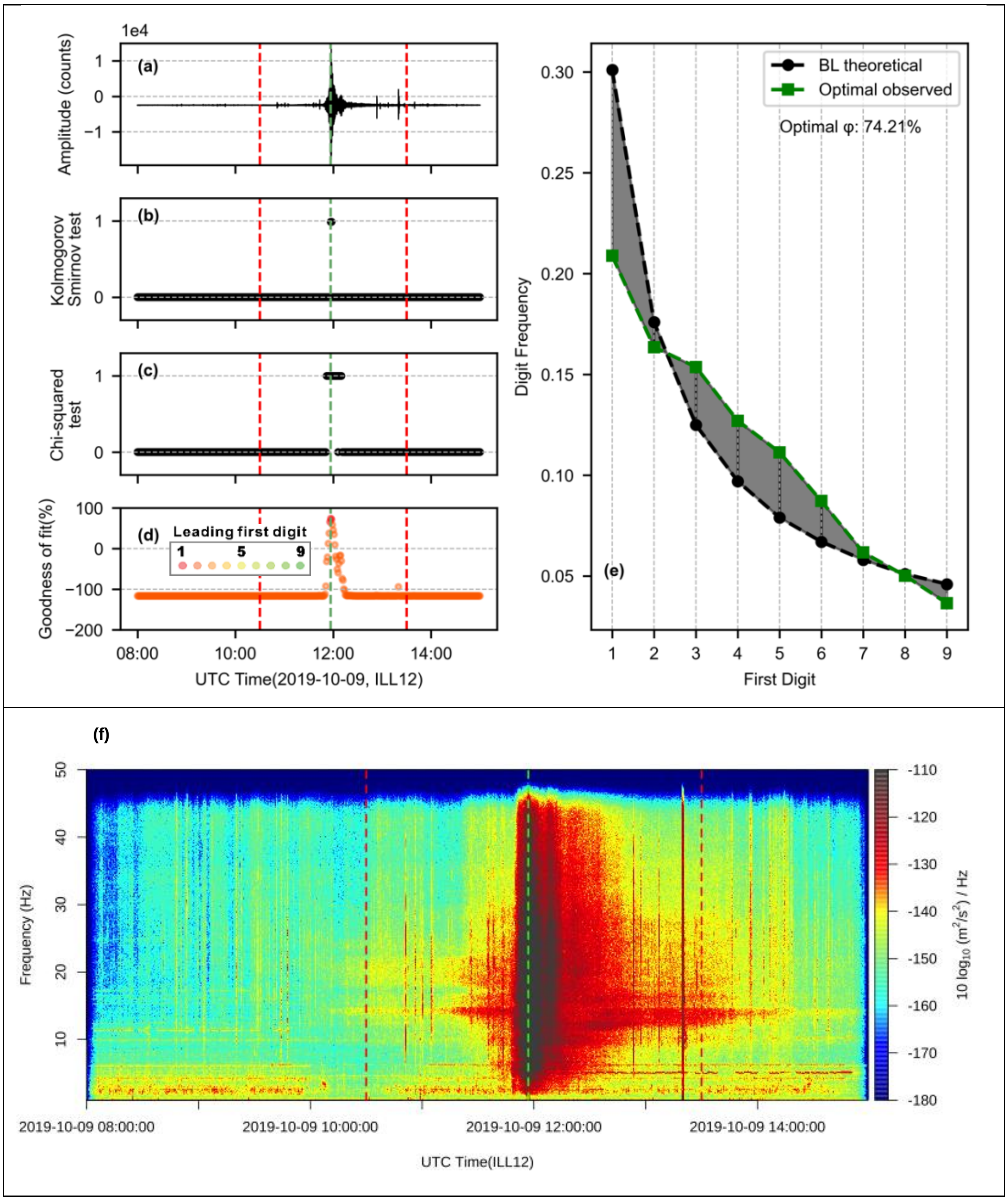


Figure S5. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (2019-10-09, validation events, WSL label, ILL12 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

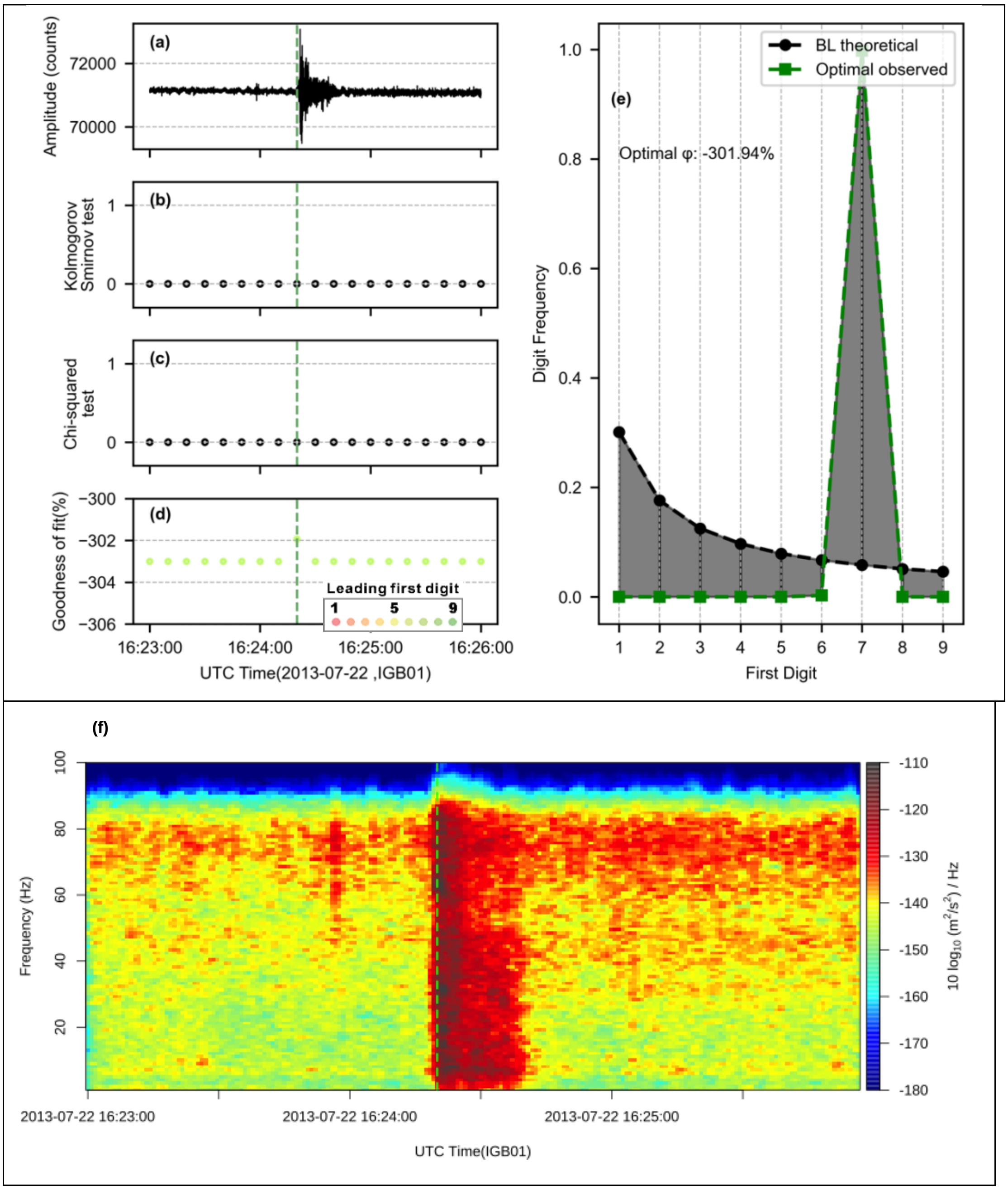


Figure S6. Raw waveform (a) and spectrogram (f) generated by the **Rockfall 1** (2013-07-22 IGB01 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e).

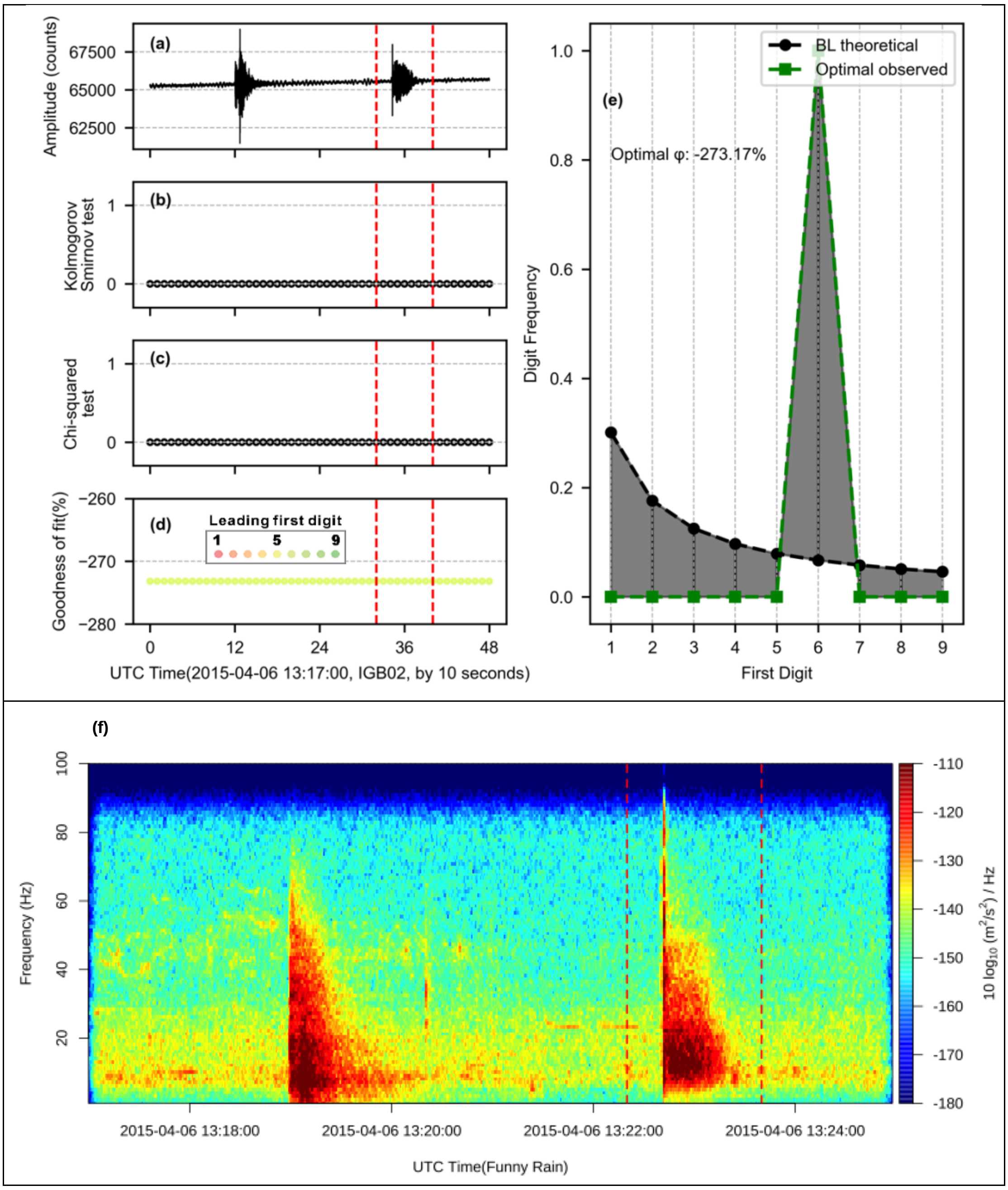


Figure S7. Raw waveform (a) and spectrogram (f) generated by the **Rockfall 2** (2015-04-06, Funny Rain station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The first event was a small long-distance earthquake. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

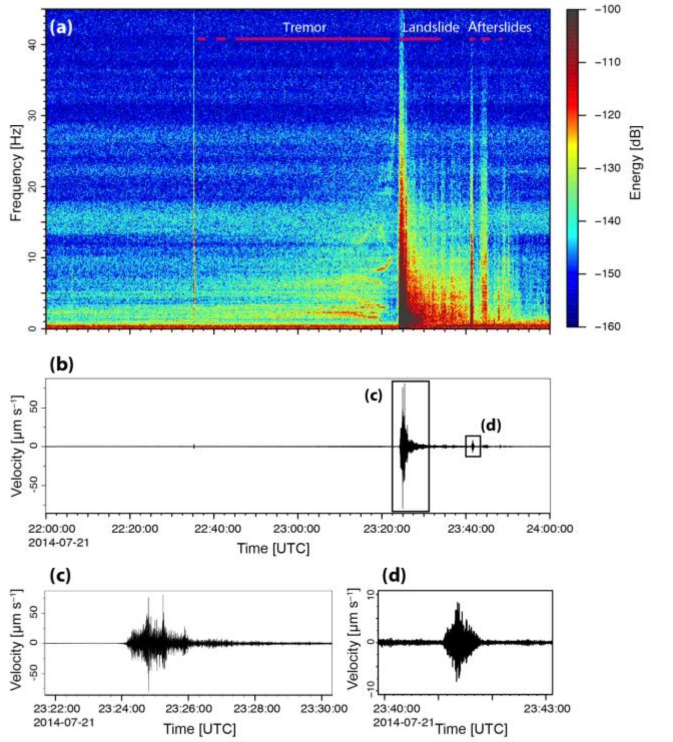


Figure S8-1. Spectrogram (a, unfiltered) and waveform (b, filtered between 1 and 45 Hz show precursory tremor and the after-slides at the Askja caldera on 21 July 2014. (c) Close-up of the landslide and (d) of one after-slide waveform. MOFO station, east component (Schöpa et al., 2018).

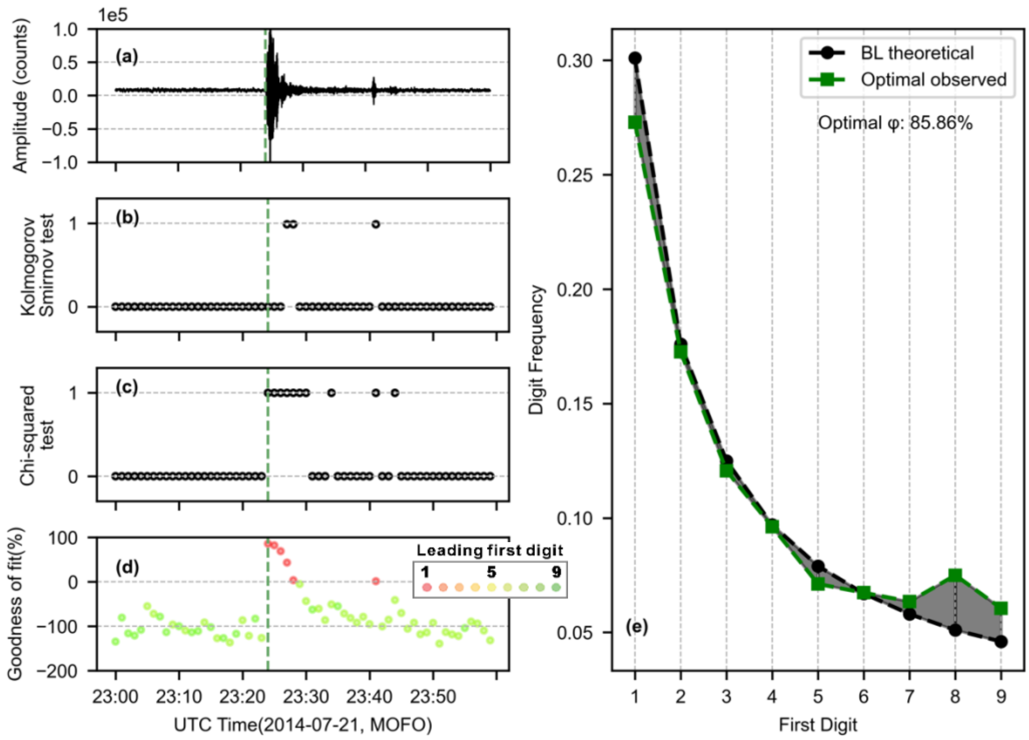


Figure S8-2. Raw waveform (a) and spectrogram (f) generated by the **landslide** (2014-07-21, MOFO station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e).

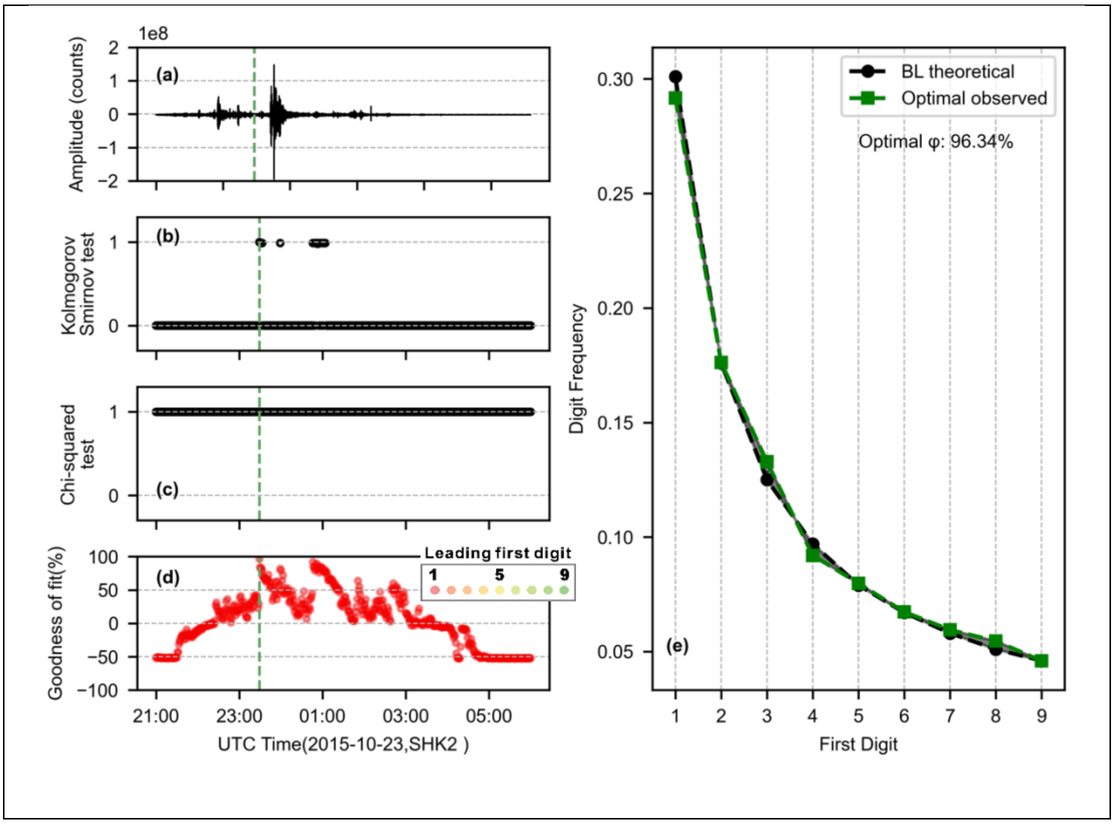


Figure S9. Raw waveform (a) generated by the a **hurricane-induced lahars** (2015-10-24 01:00, (Capra et al., 2018)). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e). *No instrument response information is available for spectrogram*.

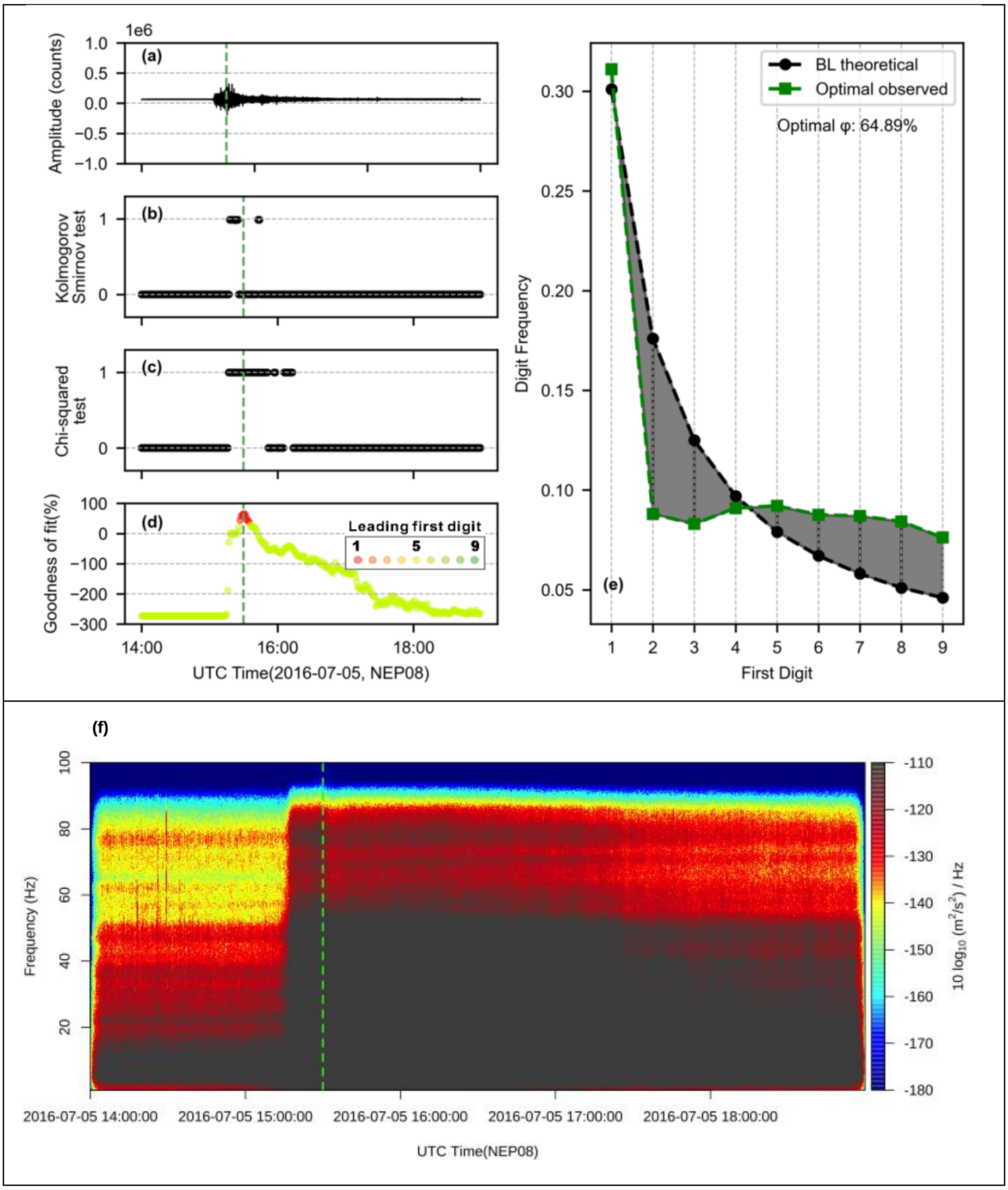


Figure S10. Raw waveform (a) and spectrogram (f) generated by a **glacial lake outburst flood** (2016-07-05, Hindi (NEP08) station, Cook et al., 2018). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e).

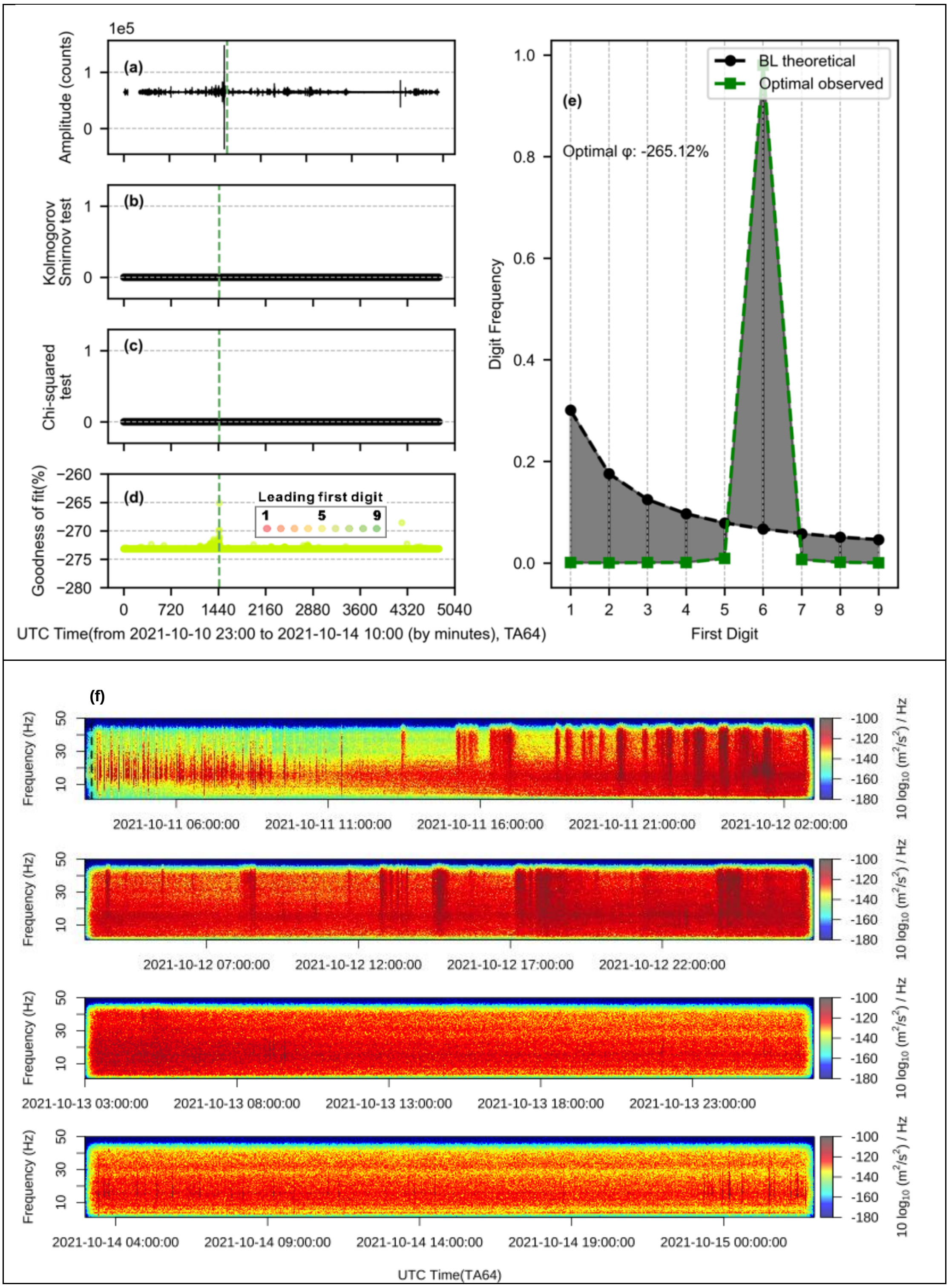


Figure S11. Raw waveform (a) and spectrogram (f) generated by a **bedload transport** (from 2021-10-10 to 2021-10-14, TA64 station). (b) and (c) results from Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First digit distribution (e) of BL theoretical and observed optimal periods. The black dashed vertical line is the time for an optimal of the goodness of fit (e).

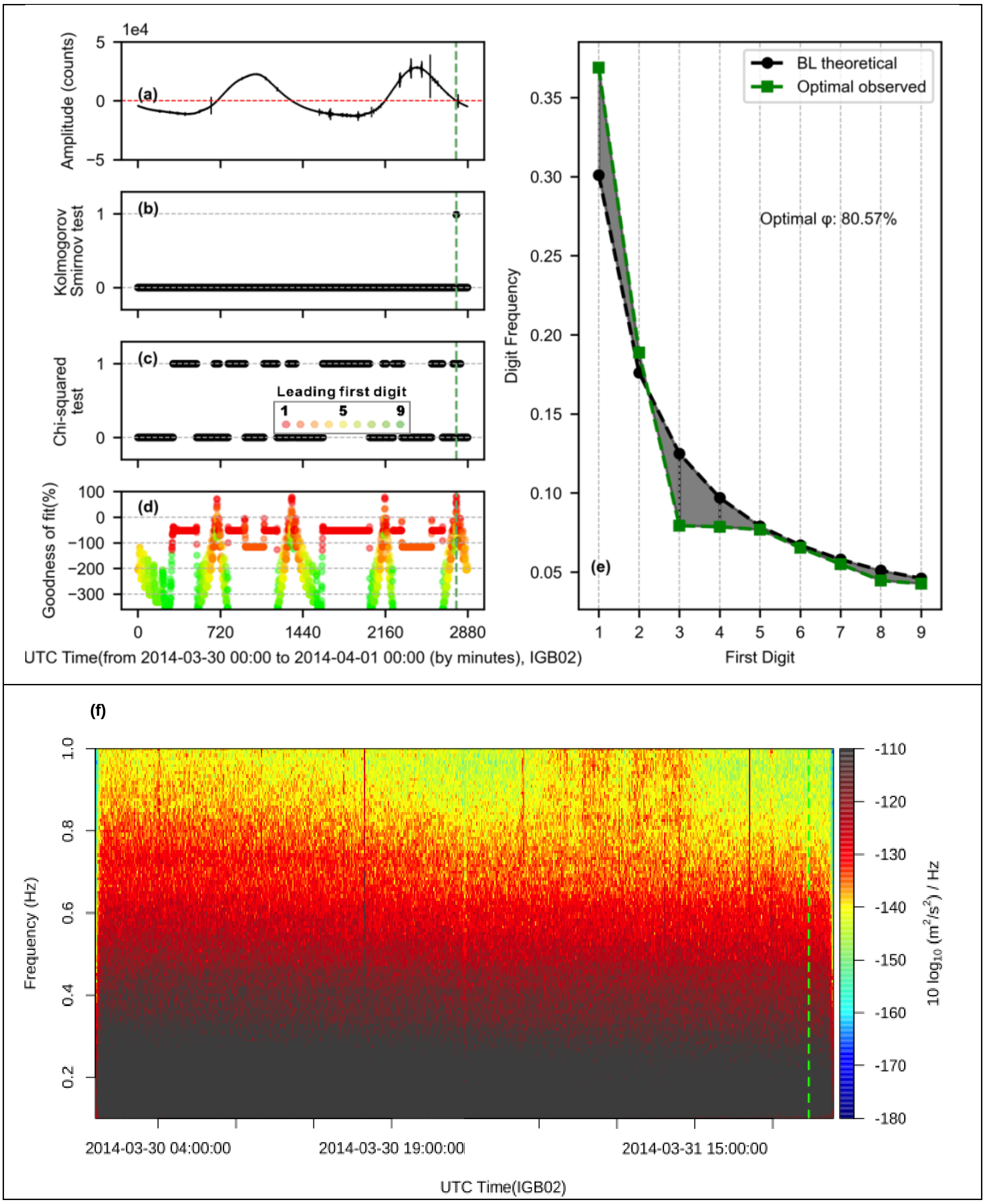


Figure S12. Raw waveform (a) and spectrogram (f) generated by **a long-period seismic signals (LP) and when LP is close to zero counts fluctuation** (from 2014-03-30 to 2014-04-01, IGB02 station, follow BL). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e). The max and min is -227 counts and 342 counts, the *iq* is 119 counts at 2014-03-31 22:22 to 22:23.

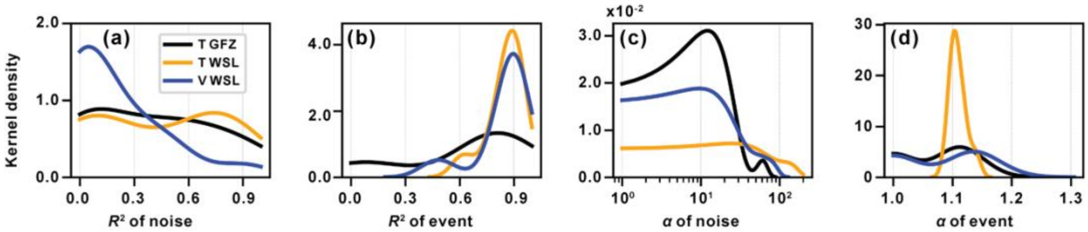


Figure S13. Kernel density for exponential fitting and power law exponent. (a) and (b) are the kernel density (bandwidth is 0.5) of R^2 of the exponential fitting. The coefficients of the exponential fitting for all events are listed in Table S5. (c) and (d) are kernel density (bandwidth is 0.5) of power law exponent α. The data displayed in (a) to (d) correspond to the optimal goodness of fit during 38 BL-followed debris flow events. T GFZ and T WSL represent training event datasets during 2013-2014 with GFZ and WSL labels, respectively. V WSL represents the validation event dataset during 2017-2019.

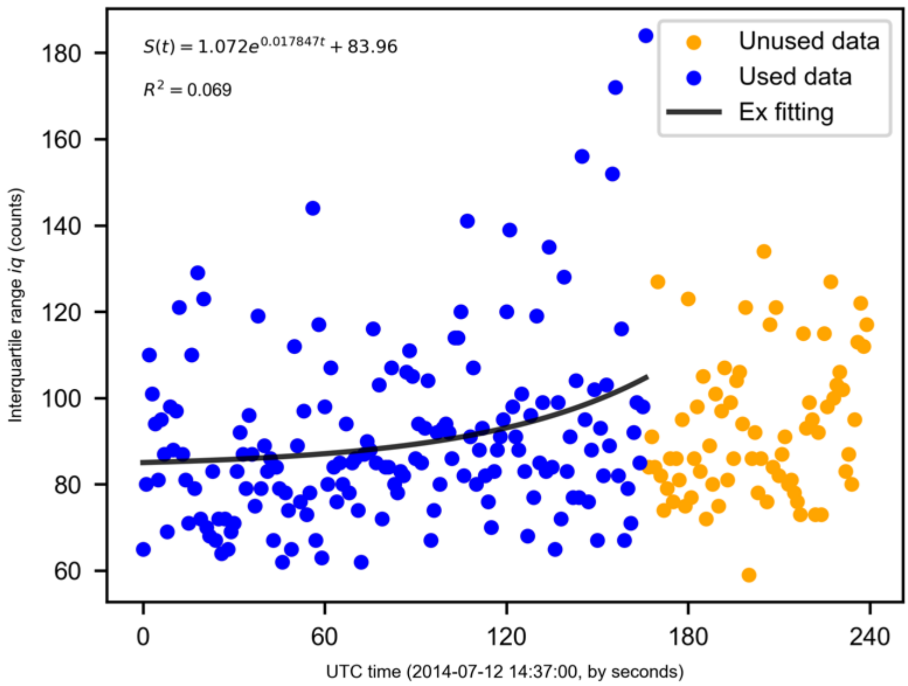


Figure S14. Exponential fitting of event 2014-07-12 debris flow noise.

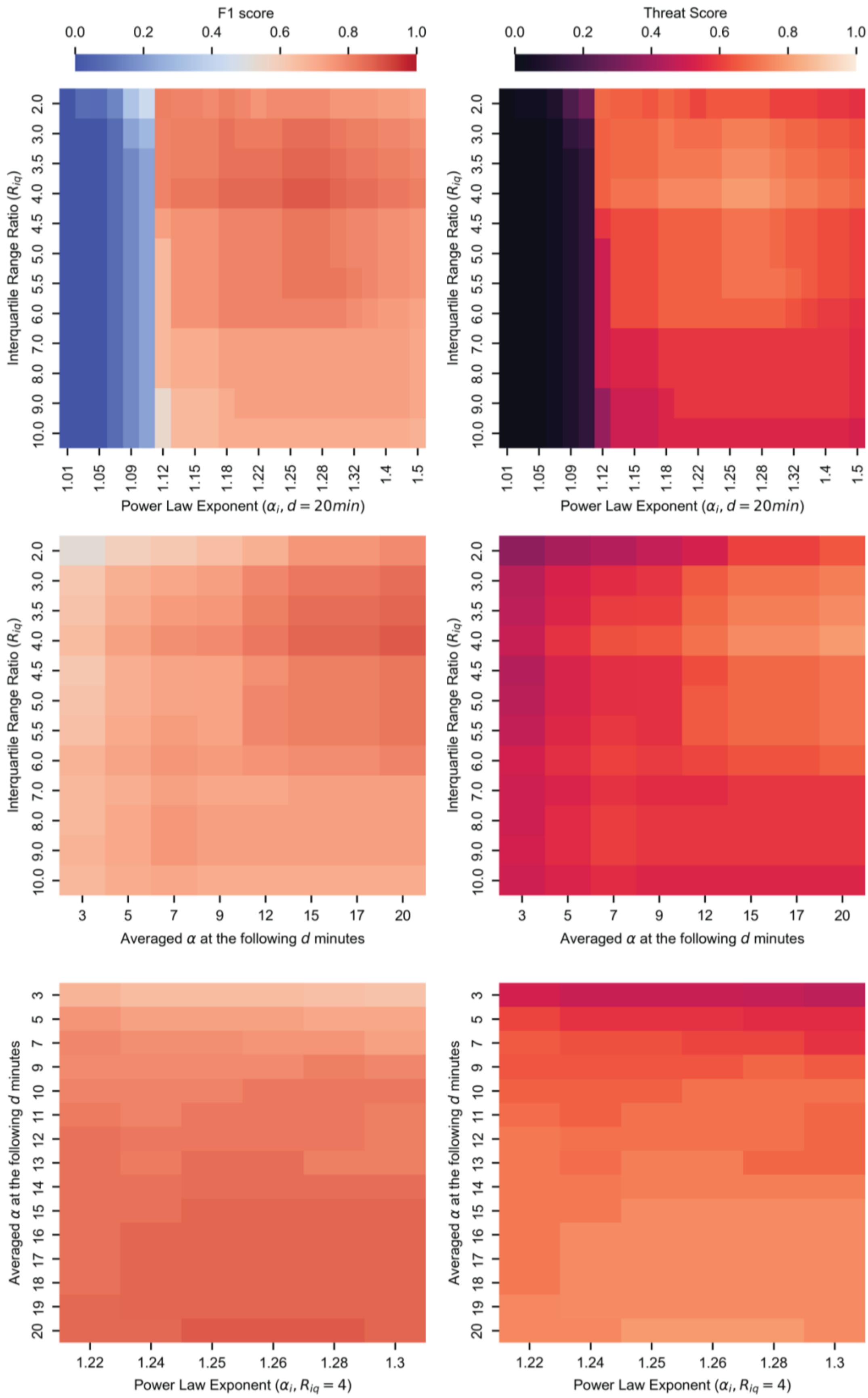


Figure S15. F1 score and Threat score for different debris-flow detector parameters.

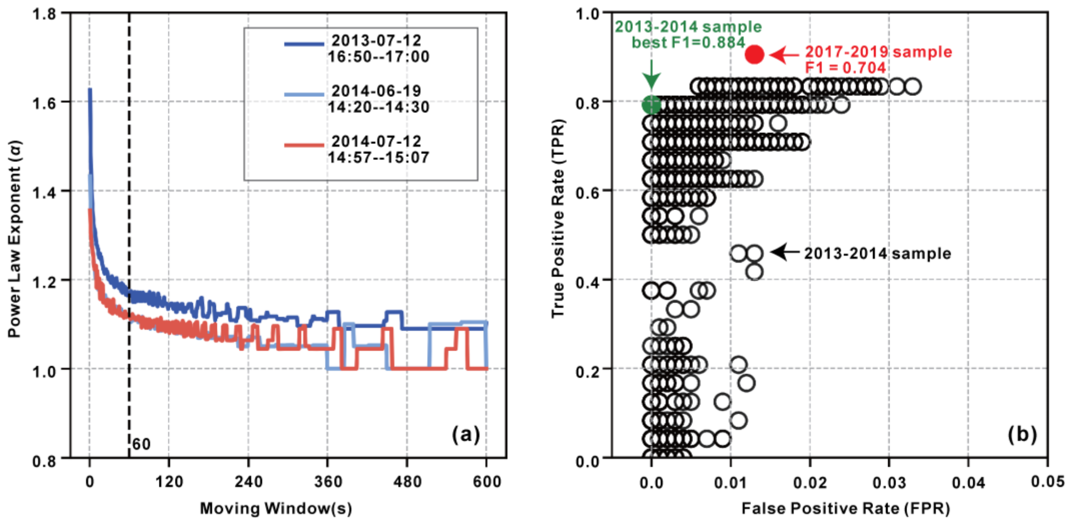


Figure S16. Variation of power law exponent with different moving window sizes and receiver operating characteristic of different detectors. (a) Averaged power law exponent α with different moving windows. The black dashed line indicates the 60 seconds window used in this work. (b) Receiver operating characteristic ROC of detectors with different parameters. The best detector (F1 0.884) obtained from the 2013-2014 training sample corresponds to FPR is 0, TPR is 0.792, and outputs of this detector for the sample from 2017-2019 FPR is 0.013, TPR is 0.905.

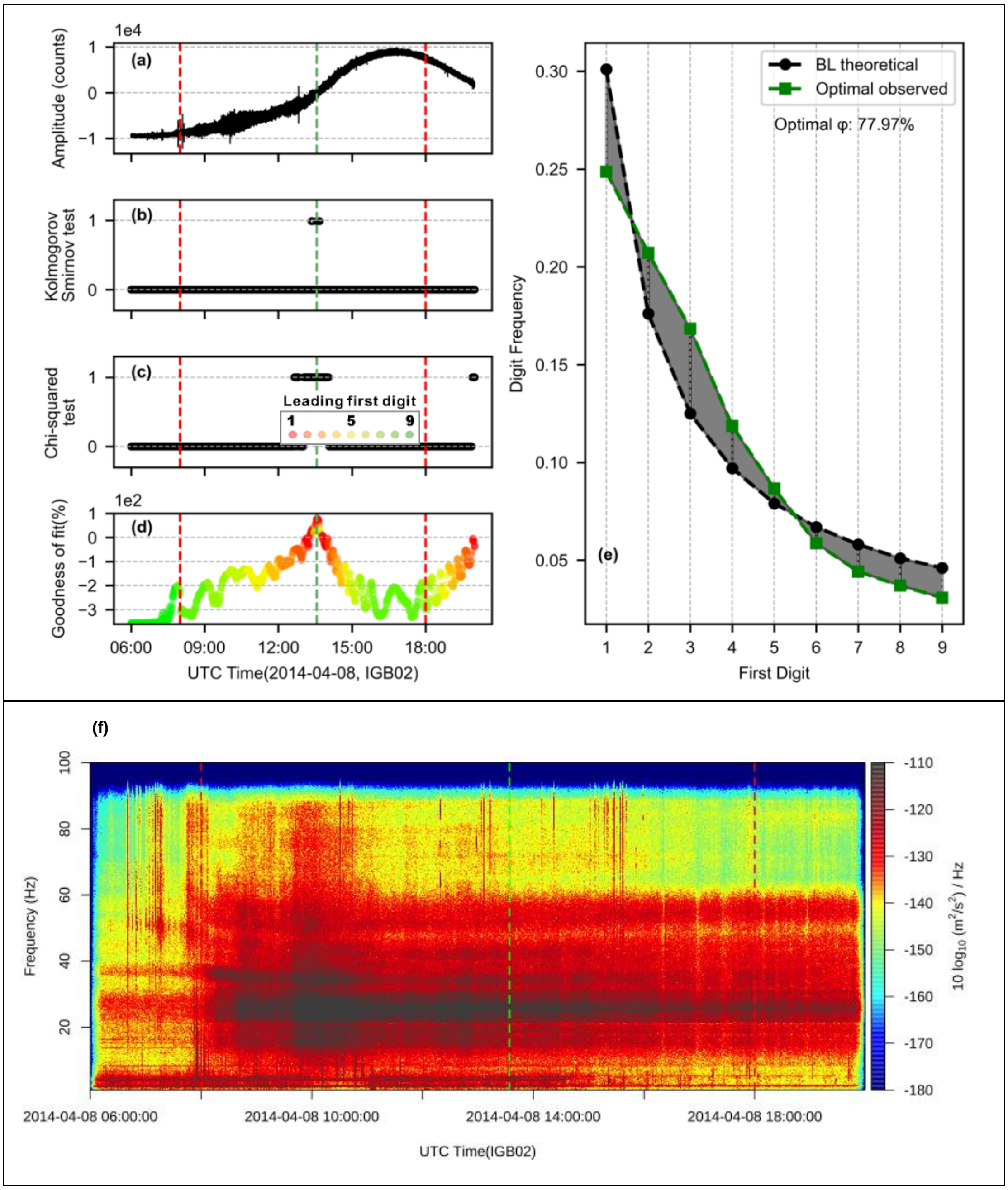


Figure S17. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-04-08, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

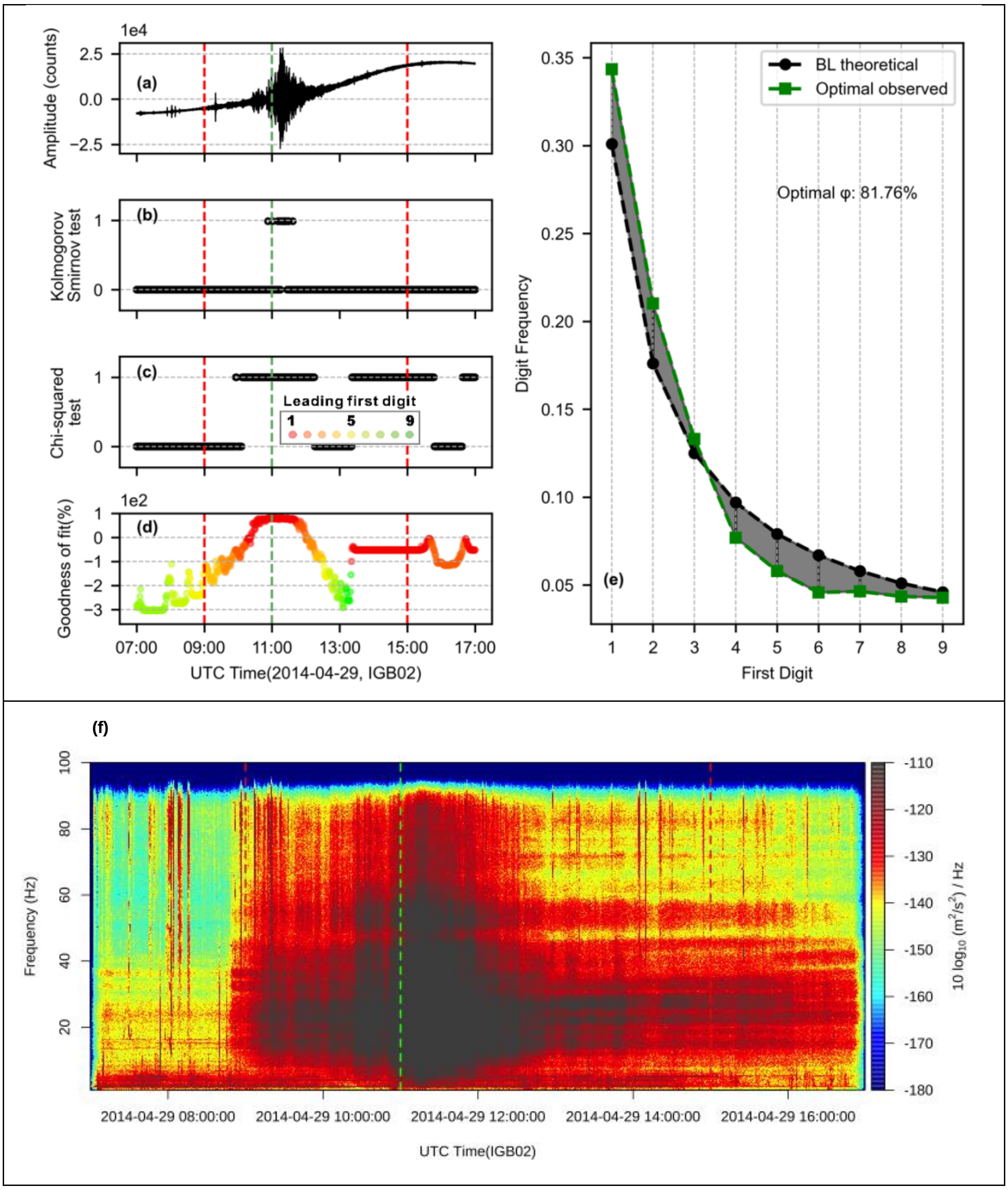


Figure S18. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-04-29, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

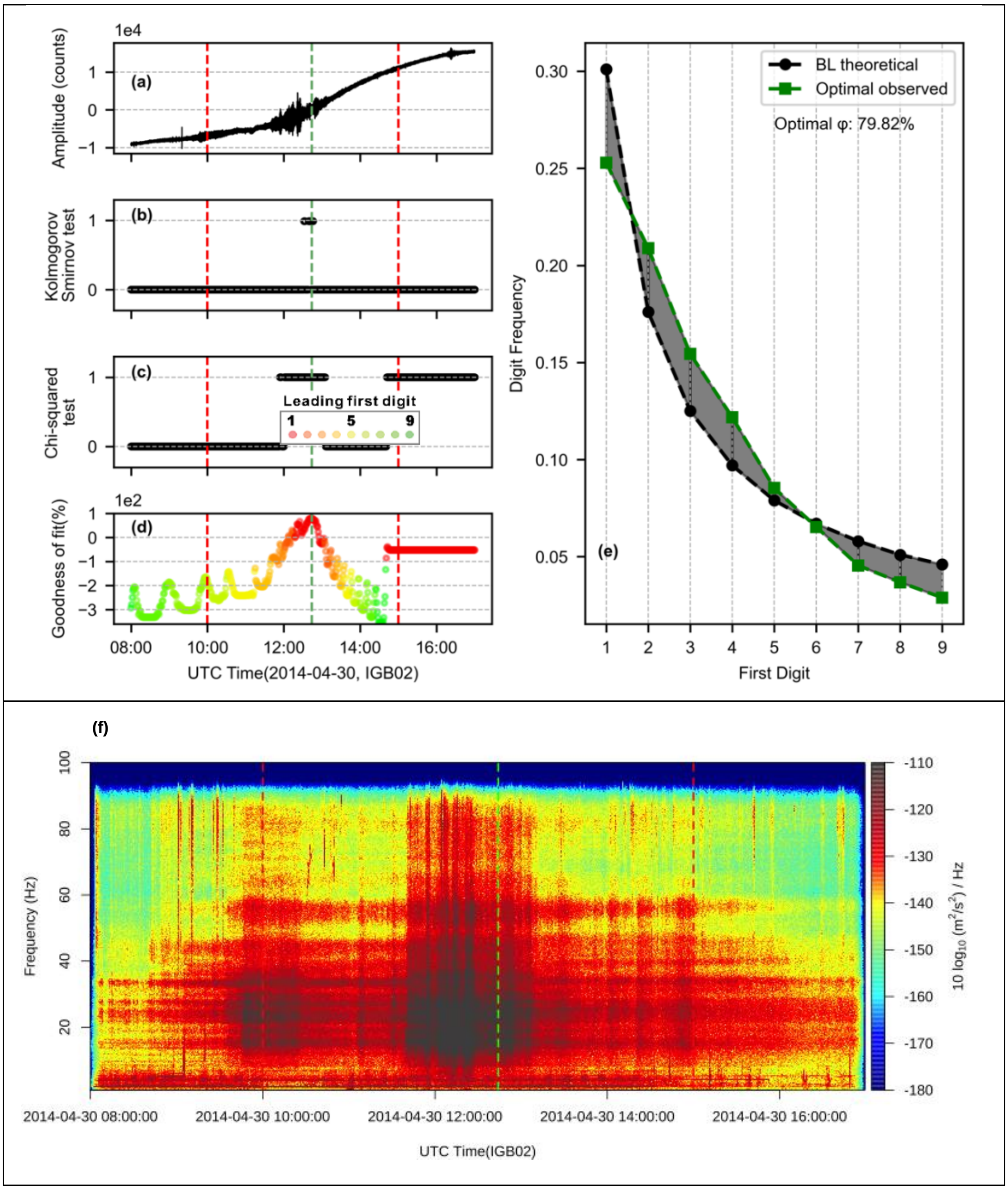


Figure S19. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-04-30, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

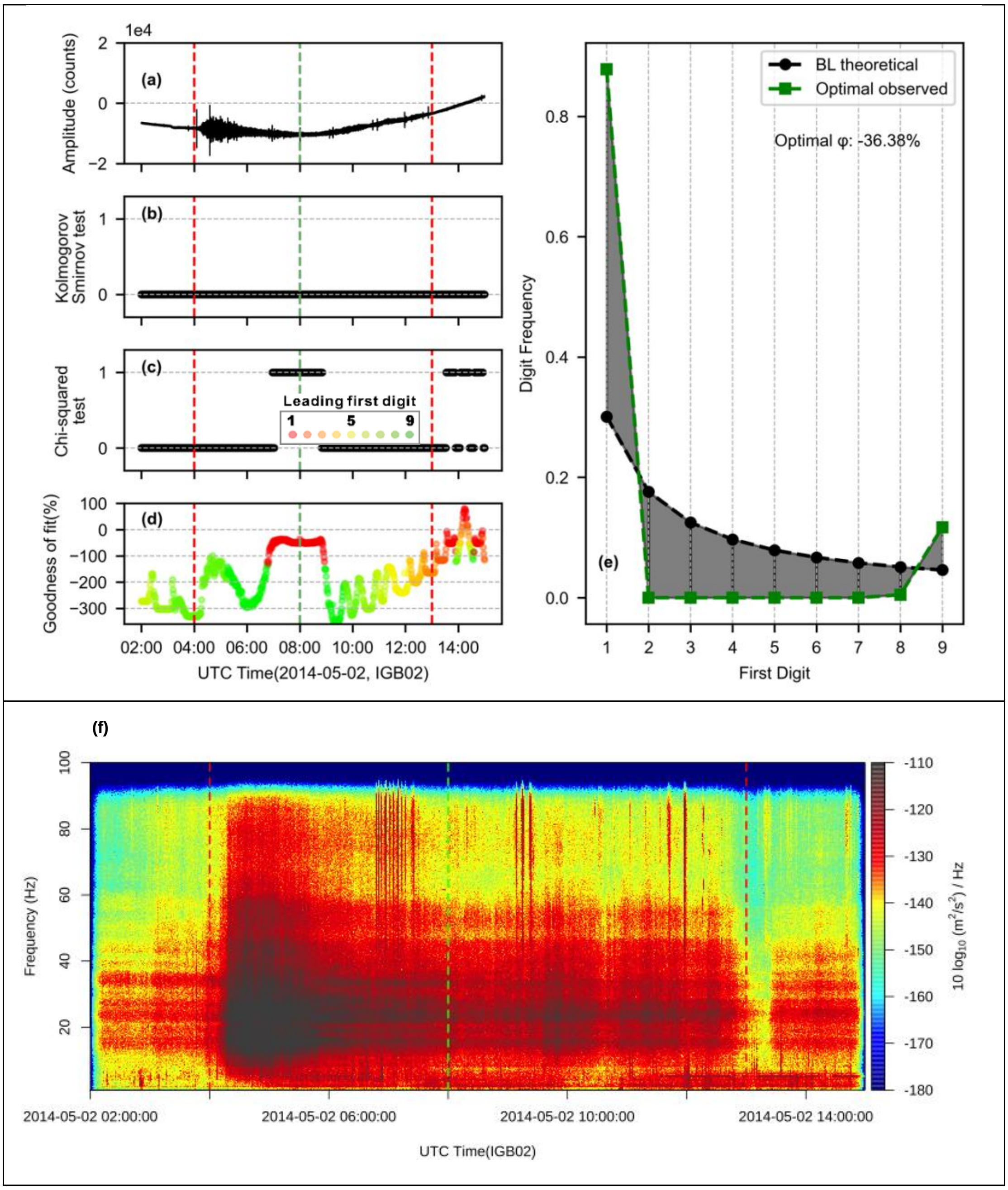


Figure S20. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-05-02, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

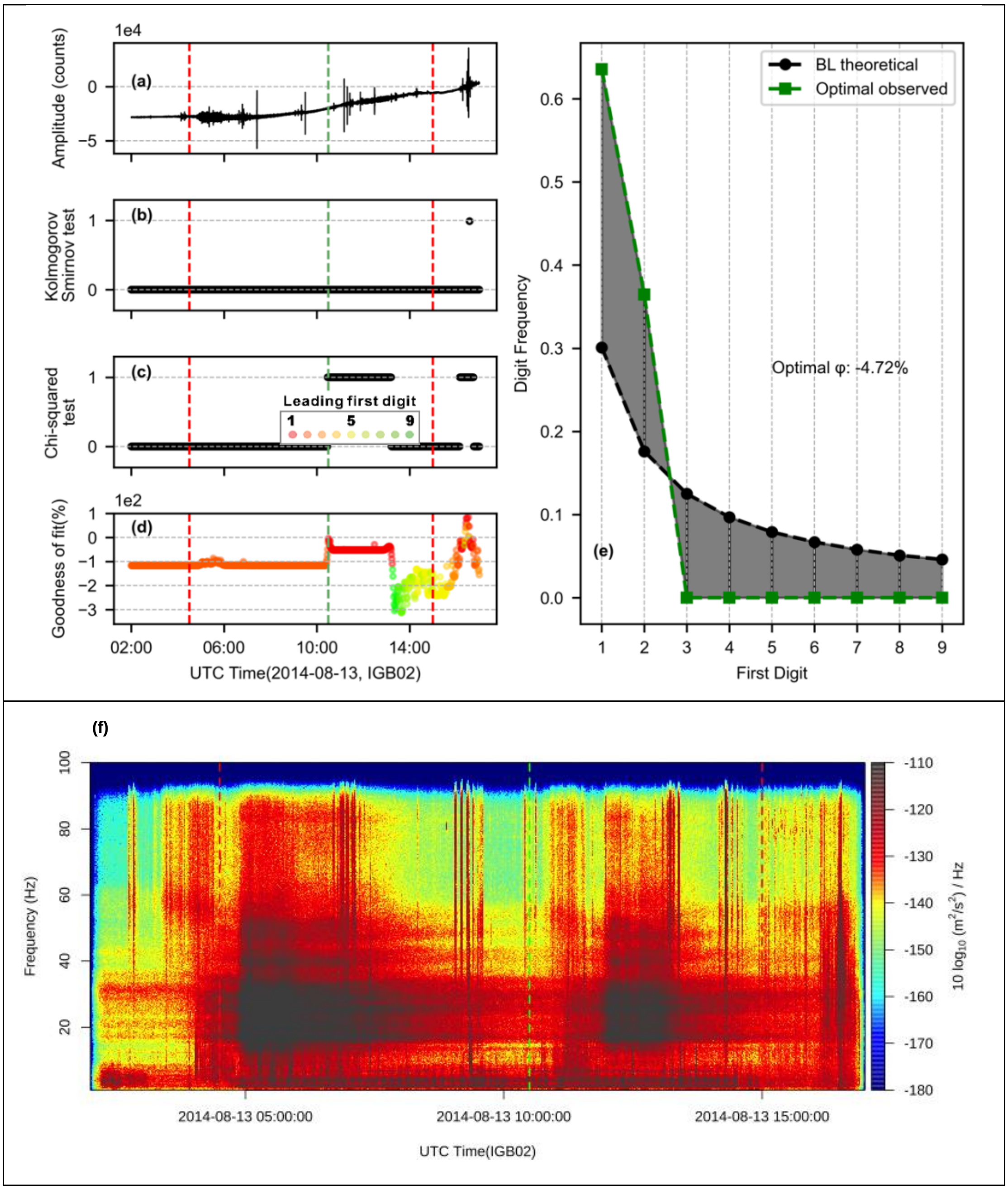


Figure S21. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-08-13, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

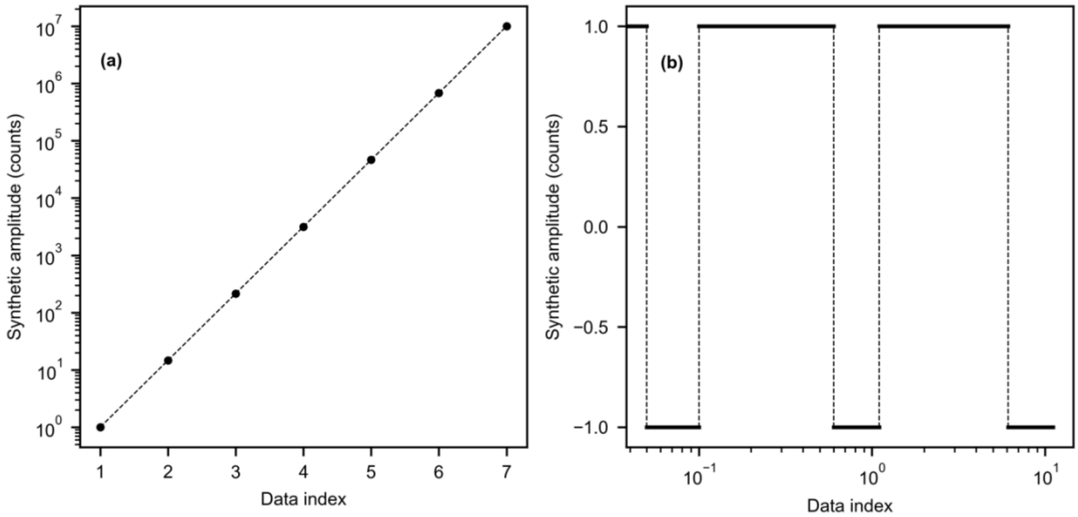


Figure S22. Synthetic amplitude dataset. (a) Datasets [1, 10, ..., 10^7] with a first-digit of 1 that span 7 orders of magnitude but do not follow BL. (b) Datasets with a first digit of 1 that has a period of 0.1 Hz (data index from 0 to 0.1), 1 Hz (data index from 0.1 to 1.1), and 10 Hz (data index from 1.1 to 11.1) but do not follow BL.

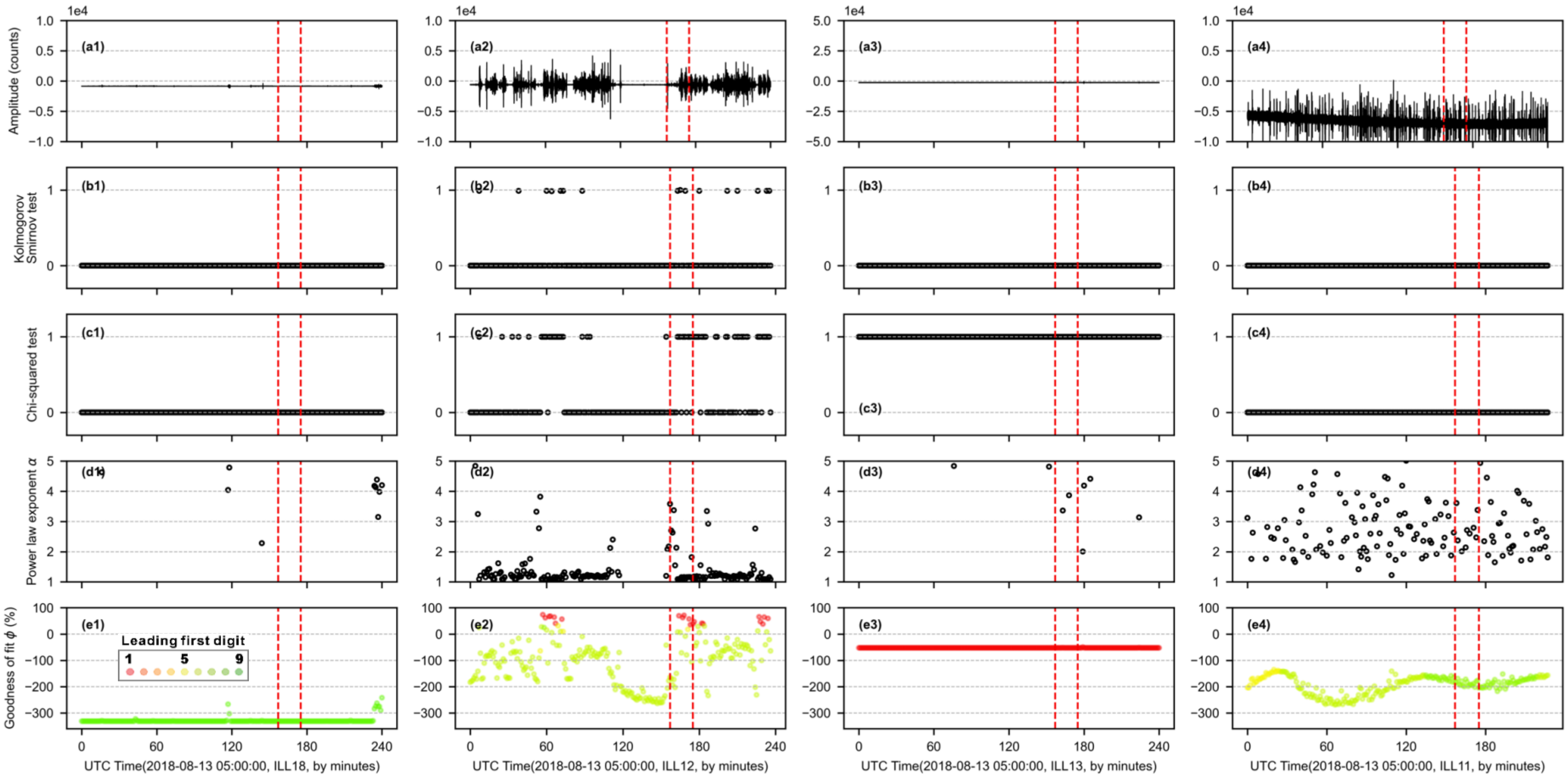


Figure S23. BL features of seismic signal generated by **Actual negative** class (2018-08-13, validation case). It was labeled as **Predicted positive** by our detector **(False positive)**. Raw waveform (row a), Chi-squared test (row b) and Kolmogorov-Smirnov test (row c) results of first-digit distribution (one means to follow BL), power-law exponent (row d), and goodness of fit φ (row e) of different seismic stations (column 1-4). The red dashed vertical lines are event start and time (2018-08-13 07:47:00, 08:06:00) of our detector marker based on ILL12. The upstream ILL18 and downstream ILL11 do not capture the same trend, this process could be caused by instruments noise or a local event.

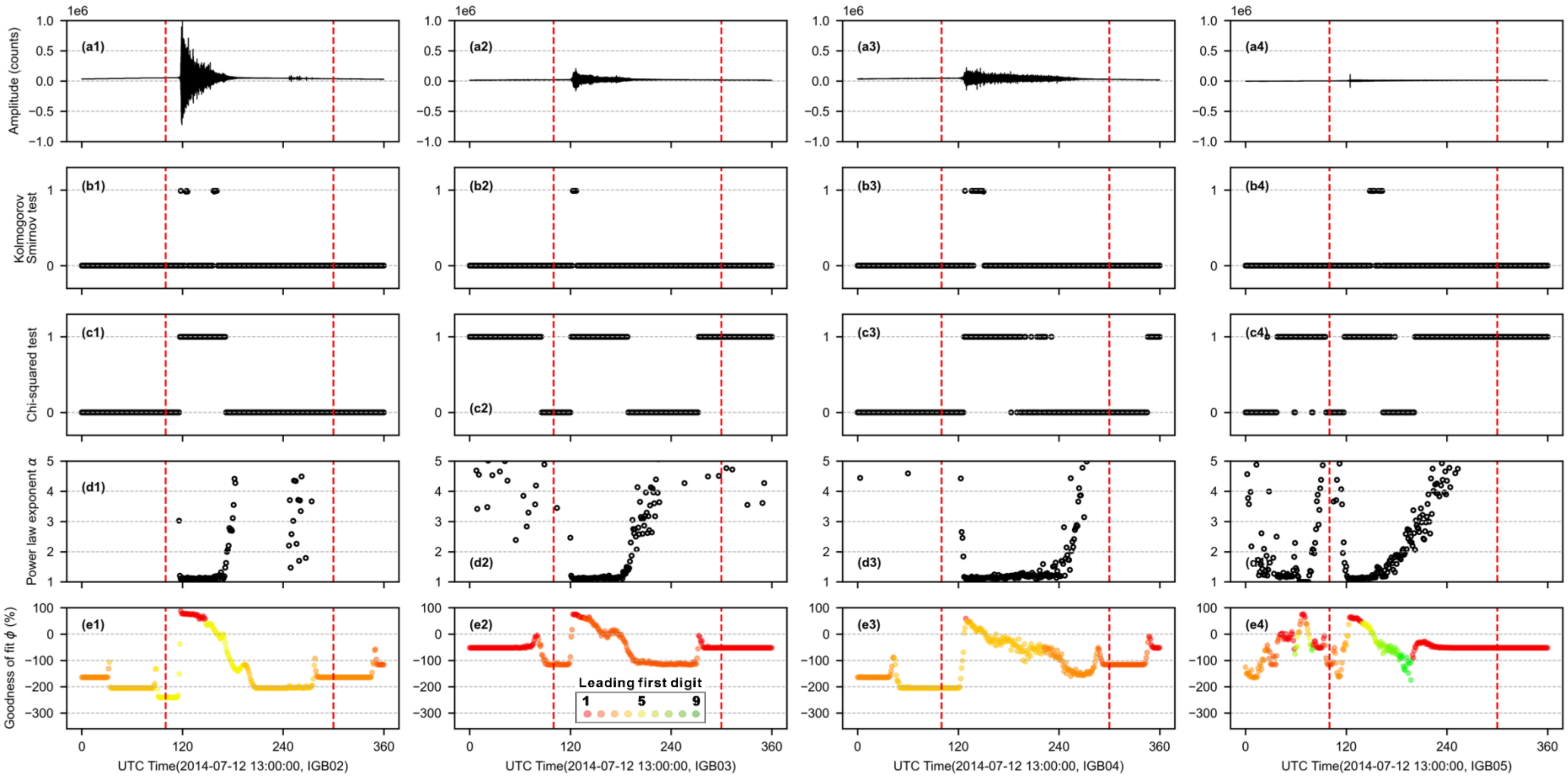


Figure S24. BL features of seismic signal generated by **Actual positive** (2014-07-12, training case). It was labeled as **Predicted positive** by our detector **(True positive)**. Raw waveform (row a), Chi-squared test (row b) and Kolmogorov-Smirnov test (row c) results of first-digit distribution (one means to follow BL), power-law exponent (row d), and goodness of fit φ (row e) of different seismic stations (column 1-4). The red dashed vertical lines are manually labeled start and end time (2014-07-12 14:40 and 18:00). IGB02 are closer to channel, and more than one station can detect this event through BL.

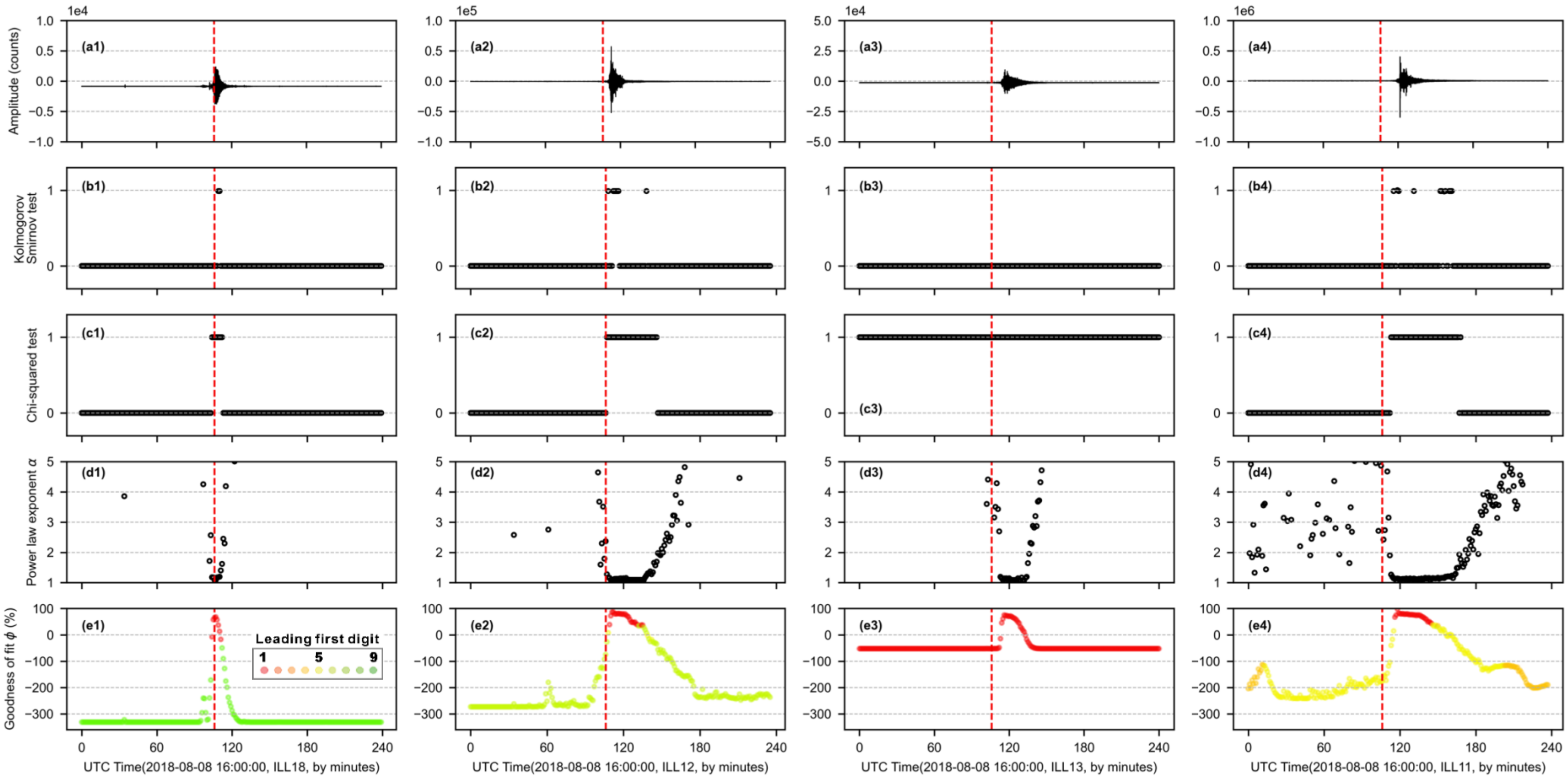


Figure S25. BL features of seismic signal generated by **Actual positive** (2018-08-08, validation case). It was labeled as **Predicted positive** by our detector **(True positive)**. Raw waveform (row a), Chi-squared test (row b) and Kolmogorov-Smirnov test (row c) results of first-digit distribution (one means to follow BL), power law exponent (row d), and goodness of fit φ (row e) of different seismic stations (column 1-4). The red dashed vertical line is debris flow front arrive at CD1 (2018-08-08 17:49:25). From upstream ILL18 to downstream ILL11, it is clearly seen that BL is captured sequentially in the time series, which indicates an event of movement along the trench.

Table S1. Seismometer Parameters

Year	Sensor	Data Logger	Signal Bandwidth	Sampling Frequency	Stations	Operated	Format
2013-2014	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	1-100 Hz	200 Hz	IGB02	GFZ[1]	SAC
2017	Lennartz LE-3D/5S	Nanometrics Centaur	1-50 Hz	100 Hz	ILL02	WSL[2]	MiniSEED
2018-2019	Lennartz LE-3D/5S	Nanometrics Centaur	1-50 Hz	100 Hz	ILL12	WSL[2]	MiniSEED

[1]GFZ, German Research Centre for Geosciences;

[2]WSL, Swiss Federal Institute for Forest, Snow and Landscape Research.

Table S2. Training Events from 2013 to 2014

N	Source	Type	Start time (UTC)	End time (UTC)	SNR	Available Station	BL
N1	WSL	F	2013-05-03 07:45	2013-05-03 08:45	1.3	01, 06, 07, 10	
N2	WSL	DF	2013-08-08 10:30	2013-08-08 12:30	1.7	01, 04, 07, 08, 10	
N3	WSL	DF	2013-08-24 13:30	2013-08-24 15:00	1.4	01, 04, 07, 08, 10	
N4	WSL	DF	2013-09-08 19:00	2013-09-08 21:00	1.7	01, 04, 07, 08, 10	
1	WSL	DF	2013-07-22 16:30	2013-07-22 18:30	4.9	01, **02**, 10	Y
2	WSL	F	2013-07-29 07:30	2013-07-29 15:00	4.2	01, **02**, 03, 07, 08, 10	Y
3	GFZ	DF	2014-04-08 08:00	2014-04-08 18:00	42.4	**02**, 03, 05, 06, 07	Y
4	GFZ	DF	2014-04-26 22:30	2014-04-27 05:00	12.9	**02**, 03, 05, 06	Y
5	GFZ	DF	2014-04-29 09:00	2014-04-29 15:00	5.9	**02**, 05, 06, 07	Y
6	GFZ	DF	2014-04-30 10:00	2014-04-30 15:00	27.0	**02**, 05, 06, 07	Y
7	GFZ	DF	2014-05-02 03:00	2014-05-02 13:00	0.1	**02**, 03, 05, 06, 07	N
8	WSL	DF	2014-05-07 15:30	2014-05-07 18:00	15.2	**02**, 03, 05, 06, 07	Y
9	GFZ	DF	2014-05-23 03:00	2014-05-23 15:00	57.0	01, **02**, 04	Y
10	GFZ	DF	2014-05-24 09:00	2014-05-24 15:00	16.1	01, **02**, 04, 05	Y
11	GFZ	DF	2014-05-27 05:30	2014-05-27 16:00	22.2	01, **02**, 03, 04, 05	Y
12	GFZ	DF	2014-06-19 14:00	2014-06-19 16:00	6.2	01, **02**, 04, 05	Y
13	GFZ	DF	2014-06-23 18:00	2014-06-23 20:00	10.1	01, **02**, 04, 05	Y
14	WSL	DF	2014-07-08 02:00	2014-07-08 14:00	10.1	01, **02**, 03, 04, 05	Y
15	WSL	DF	2014-07-12 14:40	2014-07-12 18:00	20.6	01, **02**, 03, 04, 05	Y
16	WSL	DF	2014-07-20 20:00	2014-07-21 03:00	12.7	01, **02**, 03, 04, 05	Y
17	WSL	DF	2014-07-23 23:00	2014-07-24 05:00	65.6	01, **02**, 04	Y
18	WSL	DF	2014-07-28 16:00	2014-07-28 20:50	54.3	01, **02**, 03, 04, 05	Y
19	WSL	DF	2014-07-28 20:55	2014-07-29 06:00	19.5	01, **02**, 03, 04, 05	Y

20	GFZ	DF	2014-08-02 17:30	2014-08-02 21:30	45.7	01, **02**, 03, 04, 05	Y
21	GFZ	DF	2014-08-08 18:30	2014-08-08 21:00	16.8	01, **02**, 04, 05	Y
22	GFZ	DF	2014-08-11 02:00	2014-08-11 18:00	7.7	01, **02**, 03, 04, 05	Y
23	GFZ	DF	2014-08-13 04:30	2014-08-13 15:00	0.4	01, **02**, 04, 05	N
24	WSL	DF	2014-09-08 19:30	2014-09-08 23:00	175.6	01, **02**, 04, 05	Y

N: Event number (N1-N4 were non-used events because data from IGB02 is not available). **Type**: DF is debris flow, F is flood. **Start time and End time** are labeled manually based on data from the IGB02 station. The format is *yyyy-mm-dd hh:mm* format. **SNR**, signal-to-noise ratio (N1-N4 none-used event by IGB10, 1-24 used event by IGB02). **Available station**, IGB01 abbreviation is 01. The total duration of the 24 used events is approximately 166 hours. The calculating window size is 60 seconds with 12000 data points. **BL**, from the **Start time** to the **End time** of the event, if the output of any time window from Chi-squared and Kolmogorov-Smirnov tests accepts the null hypothesis, that event will be marked as Y (Yes, follow BL); otherwise, N (No, does not follow BL).

Table S3. Validation Events from 2017 to 2019

N	Type	CD1 Time (UTC)	Start Time (UTC)	End Time (UTC)	Volume (m^3)	Data Source	BL
25	DF	2017-05-19 11:41:00	2017-05-19 10:00	2017-05-19 13:00	No Data	No Data	No data
26	DF	2017-05-29 16:58:31	2017-05-29 15:00	2017-05-29 19:00	100000	ILL02	Y
27	DF	2017-06-03 20:23:07	2017-06-03 18:00	2017-06-03 22:30	No Data	ILL02	Y
28	DF	2017-06-03 23:27:38	2017-06-03 22:40	2017-06-04 02:00	25000	ILL02	Y
29	DF	2017-06-14 19:30:48	2017-06-14 18:00	2017-06-14 23:30	No Data	ILL02	Y
30	DF	2018-06-11 10:46:39	2018-06-11 09:00	2018-06-11 14:00	35000	ILL12	Y
31	DF	2018-06-12 18:29:16	2018-06-12 17:00	2018-06-12 22:00	No Data	ILL12	Y
32	DF	2018-07-25 16:56:40	2018-07-25 15:00	2018-07-25 20:00	<50000	ILL12	Y
33	DF	2018-08-08 17:49:25	2018-08-08 16:00	2018-08-08 20:00	<100000	ILL12	Y
34	DF	2019-06-10 17:02:51	2019-06-10 16:00	2019-06-10 20:00	3300	ILL12	N
35	DF	2019-06-10 22:01:17	2019-06-10 21:00	2019-06-11 03:00	6600	ILL12	Y
36	DF	2019-06-20 09:12:17	2019-06-20 07:00	2019-06-20 11:00	No Data	ILL12	N
37	DF	2019-06-21 19:34:42	2019-06-21 19:00	2019-06-21 22:00	83000	ILL12	Y
38	DF	2019-07-01 23:00:29	2019-07-01 22:00	2019-07-02 04:00	78000	ILL12	Y
39	DF	2019-07-02 22:09:28	2019-07-02 21:00	2019-07-02 23:30	39000	ILL12	Y
40	DF	2019-07-03 16:43:15	2019-07-03 15:00	2019-07-03 20:00	No Data	ILL12	N
41	DF	2019-07-15 03:40:21	2019-07-15 02:00	2019-07-15 06:00	16000	ILL12	N
42	DF	2019-07-26 17:33:12	2019-07-26 16:30	2019-07-26 19:30	64000	ILL12	Y
43	DF	2019-08-11 17:02:34	2019-08-11 16:00	2019-08-11 19:00	53000	ILL12	Y
44	DF	2019-08-20 16:40:59	2019-08-20 15:00	2019-08-20 18:00	13000	ILL12	Y
45	DF	2019-10-09 11:45:28	2019-10-09 10:30	2019-10-09 13:30	No Data	ILL12	Y

46	DF	2019-10-15 16:10:50	2019-10-15 15:00	2019-10-15 23:00	No Data	ILL12	N

N: Event number. **CD1** time is the arrival time at CD1 and **Volume** is the integrated sum of discharge over the entire debris-flow wave (Chmiel et al., 2021). **The start time and End time** are labeled manually from ILL02 or ILL12 station. The format is *yyyy-mm-dd hh:mm* format. The calculating window size is 60 seconds with 6000 data points. **BL**, same as training events, Y (Yes, follow BL); N (No, does not follow BL).

Table S4. Signals Generated by Different Mass Movements and Fluvial Processes

N	Type	Location	Date (UTC)	Station	Sensor	Data Logger	Sampling Frequency	Window Size	Ref	BL
47	Rockfall 1	Illgraben, Switzerland	2013-07-22 16:24	IGB01	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	200 Hz	10 s	Burtin et al., 2016	N
48	Rockfall 2 Event 30 in ref	Lauterbrunnen, Switzerland	2015-04-06 13:23	Funny Rain-TC120s-Cube3extBOB	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	200 Hz	10 s	Dietze, Mohadjer, et al., 2017; Dietze, Turowski, et al., 2017	N
49	Landslide	Askja, Iceland	2014-07-21 23:24	MOFO	Güralp CMG-6TD	Nanometrics Taurus	100 Hz	60 s	Schöpa et al., 2018	Y
50	Hurricane-induced Lahar	Volcán de Colima, Meixco	2015-10-24 01:00	SHK2	Sercel SG-10	No info.	250 Hz	60 s	Capra et al., 2018	Y
51	Glacial-lake-outburst flood	Bhotekoshi, Nepal	2016-07-05 13:30	Hindi (NEP08)	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	200 Hz	60 s	Cook et al., 2018	Y
52	Bedload transport	Liwu, Hualien	2021-10-12 01:05	TA64	Sensor Nederland PE-6/B	Omnirecs Cube 3ext	100 Hz	60 s		N

BL, from the **Start time** to the **End time** of the event, if the output of any time window from Chi-squared and Kolmogorov-Smirnov tests accepts the null hypothesis, that event will be marked as Y (Yes, follow BL); otherwise, N (No, does not follow BL).

Table S5. Coefficients of the Exponential Fitting

N	$\varphi_{optimal}$	Time (UTC)	Data Length	R^2	*a*	*b*	*c*	*α*	BL
1	36.62	2013-07-22 16:52:00	2	0.767	8.776E+02	2.885E-02	2.602E+04	1.10	Y
2	20.60	2013-07-29 07:35:00	5	0.899	2.600E+01	2.209E-02	1.061E+03	1.14	Y
3	77.97	2014-04-08 13:34:00	5	0.246	-2.902E+05	-1.319E-05	2.905E+05	1.00	Y
4	81.90	2014-04-26 23:58:00	1	0.771	4.071E+00	6.973E-02	7.531E+03	1.11	Y
5	81.76	2014-04-29 11:00:00	4	0.808	8.628E+01	1.649E-02	1.091E+03	1.00	Y
6	79.82	2014-04-30 12:44:00	5	0.052	7.613E-02	7.834E-02	4.425E+02	1.00	Y
7	-36.38	2014-05-02 08:00:00	4	0.771	2.639E-08	8.310E-02	1.980E+02	5.22	N
8	80.78	2014-05-07 15:53:00	5	0.940	1.690E+02	1.626E-02	-3.612E+02	1.12	Y
9	84.18	2014-05-23 04:42:00	5	0.920	7.927E+01	2.083E-02	1.217E+03	1.11	Y
10	84.98	2014-05-24 09:36:00	3	0.171	3.535E-54	5.886E-01	1.561E+04	1.11	Y
11	78.36	2014-05-27 13:01:00	2	0.747	9.090E-21	1.262E+01	4.595E+02	1.00	Y
12	70.16	2014-06-19 14:25:00	1	0.692	3.480E+02	5.768E-02	2.879E+04	1.00	Y
13	74.50	2014-06-23 18:12:00	3	0.736	7.486E+02	1.854E-02	1.370E+04	1.12	Y
14	80.25	2014-07-08 10:13:00	2	0.618	6.075E+07	3.890E-06	-6.072E+07	1.10	Y
15	87.97	2014-07-12 14:58:00	5	0.944	2.274E+00	3.252E-02	1.494E+03	1.10	Y
16	79.11	2014-07-20 20:45:00	5	0.905	3.962E+03	5.455E-03	2.568E+03	1.12	Y
17	93.10	2014-07-23 23:03:00	5	0.868	1.573E-01	3.858E-02	1.902E+03	1.10	Y
18	87.46	2014-07-28 17:09:00	1	0.838	1.775E+00	9.960E-02	6.508E+04	1.09	Y
19	80.08	2014-07-29 01:00:00	2	0.904	3.577E+02	3.034E-02	2.800E+04	1.10	Y
20	82.17	2014-08-02 19:17:00	5	0.882	4.056E+02	2.646E-02	6.460E+03	1.10	Y
21	79.50	2014-08-08 19:18:00	1	0.832	3.044E+00	8.648E-02	2.808E+04	1.10	Y
22	82.95	2014-08-11 17:02:00	2	0.930	3.404E-13	9.804E-01	7.724E+02	1.15	Y
23	-4.72	2014-08-13 10:30:00	3	0.199	4.305E-14	4.446E-01	2.564E+02	32.77	N
24	89.14	2014-09-08 19:58:00	5	0.842	9.514E-01	3.641E-02	3.316E+03	1.11	Y
25	No data	2017-05-19							
26	92.75	2017-05-29 17:00:00	5	0.894	1.125E-02	3.508E-02	1.322E+02	1.00	Y
27	80.60	2017-06-03 20:30:00	5	0.794	6.382E+01	5.378E-03	2.322E+01	1.00	Y
28	86.79	2017-06-03 23:30:00	5	0.814	3.505E-01	2.558E-02	1.594E+02	1.14	Y

29	88.78	2017-06-14 19:32:00	5	0.947	1.507E-01	2.829E-02	1.659E+02	1.14	Y
30	77.82	2018-06-11 11:00:00	5	0.924	1.257E+02	6.730E-03	-1.350E+01	1.00	Y
31	73.20	2018-06-12 18:46:00	5	0.439	4.804E+05	1.561E-06	-4.800E+05	1.00	Y
32	83.37	2018-07-25 16:59:00	5	0.930	6.260E+00	1.854E-02	8.504E+01	1.00	Y
33	87.17	2018-08-08 17:52:00	5	0.910	7.524E+00	2.069E-02	2.320E+02	1.00	Y
34	68.99	2019-06-10 17:30:00	3	0.290	2.464E+01	1.999E-02	1.725E+03	1.13	N
35	77.29	2019-06-10 22:54:00	5	0.918	6.911E+01	1.106E-02	6.430E+02	1.14	Y
36	72.41	2019-06-20 09:23:00	5	0.840	2.176E+03	3.061E-03	-2.148E+03	1.15	N
37	80.18	2019-06-21 19:39:00	5	0.878	6.205E+03	2.586E-03	-6.752E+03	1.13	Y
38	80.54	2019-07-01 23:35:00	4	0.527	9.053E+02	6.602E-03	2.048E+03	1.13	Y
39	75.99	2019-07-02 22:16:00	5	0.847	6.091E+06	1.954E-06	-6.091E+06	1.13	Y
40	48.59	2019-07-03 17:08:00	3	0.511	6.445E+00	2.704E-02	9.103E+02	1.00	N
41	67.80	2019-07-15 03:49:00	5	0.854	9.492E+05	9.374E-06	-9.491E+05	1.13	N
42	81.66	2019-07-26 17:38:00	5	0.905	1.367E+02	1.071E-02	-2.983E+01	1.13	Y
43	87.39	2019-08-11 17:04:00	5	0.899	8.506E+00	2.044E-02	1.775E+02	1.12	Y
44	79.34	2019-08-20 16:49:00	5	0.894	3.303E+02	9.328E-03	2.485E+02	1.18	Y
45	74.21	2019-10-09 11:57:00	5	0.915	1.345E+02	1.019E-02	5.004E+02	1.00	Y
46	-24.42	2019-10-15 16:19:00	5	0.888	1.275E+02	6.050E-03	1.586E+01	1.22	N

N: Event number. $\boldsymbol{\varphi_{optimal}}$ is the optimal goodness of fit between manually labeled start and end time (unit by %). **Time** corresponds to the moment when $\boldsymbol{\varphi}$ is at its optimal. **Data length** is the length of the data used to fit the exponential function (before φ_{optima}, in minutes). $\mathbf{R^2}$ is the coefficient of determination of exponential fitting. ***a***, ***b***, and ***c*** are the coefficients of the exponential fitting in $S(t) = a \times e^{bt} + c$ (Text S3). $\boldsymbol{\alpha}$ is the power law exponent at optimal goodness of fit time.

Code S1.

All codes are available in *https://github.com/Nedasd/Benfords-law-in-environmental-seismology.git*

Reference

Belli, G., Walter, F., McArdell, B., Gheri, D., & Marchetti, E. (2022). Infrasonic and Seismic Analysis of Debris-Flow Events at Illgraben (Switzerland): Relating Signal Features to Flow Parameters and to the Seismo-Acoustic Source Mechanism. *Journal of Geophysical Research: Earth Surface*, *127*(6), 1–20. https://doi.org/10.1029/2021jf006576

Burtin, A., Hovius, N., McArdell, B. W., Turowski, J. M., & Vergne, J. (2014). Seismic constraints on dynamic links between geomorphic processes and routing of sediment in a steep mountain catchment. *Earth Surface Dynamics*, *2*(1), 21–33. https://doi.org/10.5194/esurf-2-21-2014

Burtin, A., Hovius, N., & Turowski, J. M. (2016). Seismic monitoring of torrential and fluvial processes. *Earth Surface Dynamics*, *4*(2), 285–307. https://doi.org/10.5194/esurf-4-285-2016

Capra, L., Coviello, V., Borselli, L., Márquez-Ramírez, V. H., & Arámbula-Mendoza, R. (2018). Hydrological control of large hurricane-induced lahars: Evidence from rainfall-runoff modeling, seismic and video monitoring. *Natural Hazards and Earth System Sciences*, *18*(3), 781–794. https://doi.org/10.5194/nhess-18-781-2018

Chmiel, M., Walter, F., Wenner, M., Zhang, Z., McArdell, B. W., & Hibert, C. (2021). Machine Learning Improves Debris Flow Warning. *Geophysical Research Letters*, *48*(3), 1–11. https://doi.org/10.1029/2020GL090874

Cook, K. L., Andermann, C., Gimbert, F., Adhikari, B. R., & Hovius, N. (2018). Glacial lake outburst floods as drivers of fluvial erosion in the Himalaya. *Science*, *362*(6410), 53–57. https://doi.org/10.1126/science.aat4981

Dietze, M. (2018). The R package "eseis"-a software toolbox for environmental seismology. *Earth Surface Dynamics*, *6*(3), 669–686. https://doi.org/10.5194/esurf-6-669-2018

Dietze, M., Mohadjer, S., Turowski, J. M., Ehlers, T. A., & Hovius, N. (2017). Seismic monitoring of small alpine rockfalls-validity, precision and limitations. *Earth Surface Dynamics*, *5*(4), 653–668. https://doi.org/10.5194/esurf-5-653-2017

Dietze, M., Turowski, J. M., Cook, K. L., & Hovius, N. (2017). Spatiotemporal patterns, triggers and anatomies of seismically detected rockfalls. *Earth Surface Dynamics*, *5*(4), 757–779. https://doi.org/10.5194/esurf-5-757-2017

Schöpa, A., Chao, W. A., Lipovsky, B. P., Hovius, N., White, R. S., Green, R. G., & Turowski, J. M. (2018). Dynamics of the Askja caldera July 2014 landslide, Iceland, from seismic signal analysis: Precursor, motion and aftermath. *Earth Surface Dynamics*, *6*(2), 467–485. https://doi.org/10.5194/esurf-6-467-2018